A bibliometric analysis of research on proteomics in *Science Citation Index Expanded*

Jiang Tan · Hui-Zhen Fu · Yuh-Shan Ho

Abstract   A bibliometric analysis was conducted to evaluate the global scientific output of proteomics research in the *Science Citation Index Expanded* from 1995 to 2010. The document types, languages, journals, categories, countries, and institutions were analyzed to obtain publication patterns. Research focuses and trends were revealed by a word cluster method related to author keywords, title, abstract, and *KeyWords Plus*. Bradford's Law and the correlation between keywords and institutions were identified to look deeper into the nature works. *Proteomics* and *Journal of Proteome Research* published the most articles in proteomics research. The researchers focused on the categories of biochemical research methods, and biochemistry and molecular biology. The USA and Harvard University were the most productive country and institution, respectively, while China was the fastest-growing country due to the support by Chinese government. The distribution of author keywords provided the important clues of hot issues. Results showed that mass spectrometry and two-dimensional gel electrophoresis had been the most frequently used research methods in the past 16 years; and cancer proteomics had a strong potential in the near future. Furthermore, biologists contributed significantly to proteomics research, and were more likely to co-operate with medical scientists.

J. Tan
Longping Branch, Graduate School, Central South University, Changsha 410125, Hunan, People's Republic of China

H.-Z. Fu · Y.-S. Ho (✉)
Trend Research Centre, Asia University, No. 500, Lioufeng Road, Wufeng, Taichung County 41354, Taiwan
e-mail: ysho@asia.edu.tw

H.-Z. Fu · Y.-S. Ho
Department of Environmental Sciences, Peking University, Beijing 100871, People's Republic of China

Introduction

Proteomics was the study of the proteome which was the entire set of proteins expressed by a given cell, tissue or organism (Anderson 1998; Blackstock and Weir 1999). It boosted our understanding of systems-level cellular behavior (Pandey and Mann 2000), promised a more radical transformation of biological and medical research, and was even considered the next step in the study of biological systems after genomics and transcriptomics (Tyers and Mann 2003). The term "proteomics" was coined in 1997 (James 1997) three years after Marc Wilkins posed "proteome" in a symposium (Wilkins et al. 1996). In the earlier years, most of the proteomics research focused on the study of single-celled organisms' complete proteome for identifying new proteins, analyzing of protein activities or connecting genome and proteome, such as yeast (Shevchenko et al. 1996; Washburn et al. 2001; Gavin et al. 2002), *Salmonella typhimurium* (Qi et al. 1996), *Dictyostelium discoideum* (Yan et al. 1997), *Haemophilus influenza* (Link et al. 1997), and *Saccharomyces cerevisiae* (Garrels et al. 1997; Alm et al. 1999). With the development of proteomics, scientists gradually turned more attention to multi-celled organisms especially for human diseases (Hanash 2003). Predictive and preventative medicine (Hood et al. 2004), drug delivery (Allen and Cullis 2004), new drug development (Persidis 1998), new drug targets discovery (Muellner et al. 1998), and new biomarkers discovery (Li et al. 2002; Harris et al. 2007; Zhang et al. 2004; Rifai et al. 2006) could benefit from it. Today, proteomics is widely applied in many fields, for instance, biochemical research methods, oncology, analytical chemistry and plant sciences. After intense basic research for more than a decade, it is time to have a look at the history and current situation of proteomics research all around the world.

Bibliometrics, firstly introduced by Pritchard (1969), was an effective method which had been widely used to analyze scientific production and research trends (Wang and Ho 2011). The *Science Citation Index Expanded* (*SCI-Expanded*) from the Web of Science databases were the most important and frequently used source database for a broad review of scientific accomplishment (Bayer and Folger 1966; Braun et al. 2000). It had widely been used for the bibliometric analysis of various fields, such as stem cell (Li et al. 2009a), horizontal gene transfer (Wen et al. 2009), solid waste (Fu et al. 2010), acupuncture (Han and Ho 2011), estuary pollution (Sun et al. 2012), and photosynthesis (Yu et al. 2012). In exited bibliometric analyses, the analyzed aspects traditionally covered languages (Alfaraz and Calvino 2004), annual publication outputs (Chiu and Ho 2007), journals (Schubert et al. 1989), categories (Moed et al. 1985), and contributing countries and institutions (Schubert et al. 1989). In particular, the five indicators including total, independent, collaborative, first author, and corresponding author articles have been recently developed to compare the publication performance of countries and institutions (Ho et al. 2010; Tanaka and Ho 2011). *h*-index, introduced in 2005 (Hirsch 2005), has been a representative indicator of scientific achievement (Bajwa et al. 2013). In recent years, title words, author keywords, and *KeyWords Plus* which could provide a reasonably detailed picture of the article's subject (Garfield 1990), have been quantitatively analyzed to figure out research emphases and trends (Chiu and Ho 2007; Fu et al. 2010; Li et al. 2011).

This study aimed to use a comprehensive method of bibliometric analysis to provide a thoroughly graph of research on proteomics from 1995 to 2010. Document types, languages, categories, journals, countries/territories and institutions were traditionally identified to characterize the proteomics research. An innovative method-word cluster analysis (Mao et al. 2010) of selected topics in the combination of paper titles, author keywords, abstracts, and *KeyWords Plus* was also applied to map the global research trends. In addition, the correlation of author keywords and institutions was also revealed to obtain

information for nature works. Findings from these investigations are helpful in realizing the breadth of proteomics research and establishing further research directions.

Methods

The data were based on the online version of the *SCI-Expanded*, the Thomson Reuters Web of Science. According to *Journal Citation Reports* (*JCR*), it indexes 8,005 major journals with citation references across 174 scientific disciplines in 2010. "Proteome", "proteom", "proteomes", "proteomic", "proteomics", "proteomies", "proteomical", "proteomical-ly", "proteomique", "proteomi", "proteomica", and "proteomiec" were searched in terms of topic (including four parts: title, abstract, author keywords, and *KeyWords Plus*) within the publication year limitation from 1995 to 2010 based on *SCI-Expanded*. Articles originating from England, Scotland, Northern Ireland, and Wales were reclassified as being from the United Kingdom (UK) (Chiu and Ho 2005). Articles from Hong Kong published before 1997 were included in the China category (Chuang et al. 2011). Besides, the reported impact factor (IF) of each journal was obtained from the JCR in 2010. Contributions of different institutions and countries/territories were estimated by the affiliation of at least one author to the articles. Collaboration type was determined by the addresses of the authors; where the term "single country article" was assigned if the researchers' addresses were from the same country. The term "internationally collaborative article" was designated to those articles that were coauthored by researchers from multiple countries/territories. The term "single institution article" was assigned if the researchers' addresses were from the same institution. The term "inter-institutionally collaborative article" was assigned if authors were from different institutions. *h*-index is defined by the h of Np papers having at least h citations each and the other (Np-h) papers have ≤h citations each (Hirsch 2005).

The following discussion included two sections to determine the scientific performances and research activity trends. The first section dealt with publication patterns of document types, languages, categories, journals, institutions and countries. The other section focused on the research emphases and trends by the author keywords and word cluster analysis, as well as the correlation of keywords and institutions.

Publication patterns

Document type and language of publication

There were 43,683 publications with 16 document types indexed in the *SCI-Expanded* during the 16-year study period, which including 27,956 articles. The article, as the most popular document type, comprised 64 % of the total production and was followed distantly by meeting abstracts (5,752; 13 %), reviews (5,665; 13 %), proceedings paper articles (2,366; 5.4 %), editorial materials (1,062; 2.4 %), news items (237; 0.54 %), book chapter reviews (224; 0.51 %), corrections (178; 0.41 %), letters (161; 0.37 %). The remainder having less significance were book chapter articles (37), software reviews (25), biographical-items (6), reprints (6), database reviews (4), book reviews (3), and bibliography (1). Journal articles represented the majority of document types, therefore 27,956 articles were analyzed in further study. There were 15 languages in use. English, as the most popular language, comprises 99 % of the total articles. Some other languages that were less

used on were as follows: French (68), German (38), Japanese (33), Chinese (21), Spanish (11), Czech (8), Polish (6), Korean (5), and Russian (2). Some other languages appeared only once were Portuguese, Hungarian, Malay, multi-language with English and French, and multi-language with English and Spanish. There were 216 (0.77 %) articles with unspecified languages.

Web of Science categories and journals

Distribution of Web of Science categories and journals had been studied in research topics (Chiu and Ho 2005). Based on the classification of categories in JCR in 2010, the article output data of proteomics research was distributed in 164 Web of Science categories in science edition. The top ten productive Web of Science categories are shown in Table 1. The categories containing at least 2,000 articles were biochemical research methods (8,899; 32 %), biochemistry and molecular biology (7,999; 29 %), analytical chemistry (2,965; 11 %), and biotechnology and applied microbiology (2,310; 8.3 %). Nearly a half of journals listed in both the categories of biochemical research methods and biochemistry and molecular biology, indicating proteomics was a technology-depended subject and played an important role in biochemistry and molecular biology (Ong and Mann 2005; Peng et al. 2003).

In total, 27,956 articles were published in 2,359 journals including specialty journals and journals of other disciplines. According to Bradford's Law of Scattering (Bradford 1934), the journals were sorted in descending order in terms of number of articles, and then divided into three "zones". Zone one represents the most productive one-third of the total articles, with 11 (0.47 %) of 2,359 journals. Zone two represents the next most productive one-third of total articles, with 149 (6.3 %) of 2,359 journals, and Zone three represents the least productive one-third of total articles with 2,199 (93 %) of 2,359 journals. The number of journals was approximately $1:n:n^2$ (1:14:200), followed Bradford's law. The 11 most productive of Bradford's core journals are listed in Table 2. *Proteomics* published the most proteome articles (2,427; 8.7 %), followed by *Journal of Proteome Research* (2,203; 7.9 %), and *Molecular and Cellular Proteomics* (1,064; 3.8 %). The percentages of the top journals were not high, which indicated the breadth of article distribution in proteomics research as well as the broad interest in proteomics from various research angles (Li et al. 2011). This phenomenon also appeared in certain fields, such as *Atmospheric Environment* (8.7 %) in atmospheric simulation (Li et al. 2009b) and *Geophysical Research Letters*

Table 1 Top 10 active categories

Web of Science category	TP	%
Biochemical research methods	8,899	32
Biochemistry and molecular biology	7,999	29
Analytical chemistry	2,965	11
Biotechnology and applied microbiology	2,310	8.3
Cell biology	1,758	6.3
Microbiology	1,319	4.7
Genetics and heredity	1,267	4.5
Oncology	1,200	4.3
Biophysics	1,014	3.6
Plant sciences	904	3.2

TP: number of total articles

Table 2 The 11 core journals on proteome, including the rankings, percentages, impact factors

Journals	IF 2010	TP (%)	Web of Science categories	Rank
Proteomics	4.815	2,427 (8.7)	Biochemical research methods	13/71
			Biochemistry and molecular biology	61/286
Journal of Proteome Research	5.460	2,203 (7.9)	Biochemical research methods	9/71
Molecular and Cellular Proteomics	8.354	1,064 (3.8)	Biochemical research methods	5/71
Analytical Chemistry	5.874	790 (2.8)	Analytical chemistry	3/73
Electrophoresis	3.569	642 (2.3)	Biochemical research methods	21/71
			Analytical chemistry	12/73
Journal of Biological Chemistry	5.328	486 (1.7)	Biochemistry and molecular biology	50/286
Proceedings of the National Academy of Sciences of the United States of America	9.771	440 (1.6)	Multidisciplinary sciences	3/59
Bioinformatics	4.877	317 (1.1)	Biochemical research methods	12/71
			Biotechnology and applied microbiology	18/160
			Mathematical and computational biology	2/37
Nucleic Acids Research	7.836	310 (1.1)	Biochemistry and molecular biology	30/286
PLoS One	4.411	238 (1.5)	Biology	12/86

IF 2010: impact factor in 2010 JCR, TP: number of total articles

(3 %) in global climate change (Li et al. 2011). In addition, as regards to IF, *Nature Genetics* won the first place with the highest IF (IF = 36.377) with 23 articles, followed by *Nature* (IF = 36.104) with 71 articles, *Cell* (IF = 32.401) with 64 articles, *Science* (IF = 31.364) with 64 articles, and *Nature Biotechnology* (IF = 31.085) with 90 articles.

Countries/territories' research performances

Each author of an article has made an independent contribution to the manuscript (Coats 2009), and therefore the country and institution the author affiliated could be consider the important contributors for the evaluation of research. Publication count of countries was a reference for evaluating countries research performance in a field and has been used in many aspects of research (Li et al. 2011; Chiu and Ho 2007). The contribution of different countries was estimated by the location of the affiliation of at least one author of the published papers. Table 3 shows the top 19 productive countries with at least 300 articles ranked by the number of total articles with six indicators. Among the 27,956 articles with author address information, international collaborations articles accounted for 24 % of the total articles, while 76 % articles were from single countries. The international collaborative rate of proteomics research is higher than that in certain studies, such as 14 % biosorption technology for water treatment (Ho 2008) and acupuncture research (Han and Ho 2011), 16 % of desalination research (Tanaka and Ho 2011) and solid waste (Fu et al. 2010). Single country articles were published by 67 different countries, most of which originated from the USA (7,778; 37 %), followed by China (2,066), Germany (1,541), Japan (1,205), and the UK (1,076). The USA had the most partners accounting for 48 % of

Table 3 Characteristics of the top 19 productive countries/territories (TP ≥ 300)

| Country	TP	TPR (%)	SPR (%)	CPR (%)	FPR (%)	RPR (%)	S%	h-index
USA	11,013	1 (40)	1 (37)	1 (48)	1 (34)	1 (34)	71	209
Germany	3,069	2 (11)	3 (7.3)	2 (23)	3 (8.1)	3 (8.0)	67	117
China	2,758	3 (10)	2 (10)	5 (10)	2 (8.8)	2 (8.7)	56	67
UK	2,458	4 (8.9)	5 (5.1)	3 (21)	4 (6.0)	4 (6.0)	49	112
France	1,680	5 (6.0)	6 (4.2)	4 (12)	6 (4.3)	6 (4.3)	64	86
Japan	1,665	6 (6.0)	4 (5.7)	10 (6.9)	5 (4.9)	5 (4.9)	53	74
Canada	1,405	7 (5.1)	8 (3.5)	6 (10)	7 (3.8)	7 (3.8)	59	91
Italy	1,242	8 (4.5)	9 (3.3)	8 (8.3)	8 (3.5)	8 (3.5)	60	60
South Korea	1,062	9 (3.8)	7 (3.9)	17 (3.6)	9 (3.4)	9 (3.4)	65	44
Australia	854	10 (3.1)	10 (2.1)	12 (6.1)	11 (2.2)	11 (2.2)	52	65
Spain	838	11 (3.0)	11 (2.1)	13 (6.0)	10 (2.3)	10 (2.4)	52	50
Netherlands	835	12 (3.0)	13 (1.7)	9 (7.0)	13 (2.0)	13 (2.0)	45	70
Switzerland	825	13 (3.0)	15 (1.2)	7 (8.5)	14 (1.6)	14 (1.6)	44	77
Sweden	799	14 (2.9)	12 (1.8)	11 (6.3)	12 (2.0)	12 (2.1)	46	64
Denmark	579	15 (2.1)	16 (0.94)	14 (5.7)	16 (1.2)	17 (1.1)	44	85
Taiwan	493	16 (1.8)	14 (1.7)	22 (1.9)	15 (1.5)	15 (1.5)	40	36
Belgium	465	17 (1.7)	18 (0.86)	15 (4.3)	18 (1.1)	18 (1.1)	42	53
Austria	461	18 (1.7)	19 (0.85)	16 (4.2)	17 (1.1)	16 (1.2)	39	49
India	321	19 (1.2)	17 (0.93)	23 (1.9)	19 (0.85)	19 (0.85)	56	32

TP: number of articles, TPR: the rank of total articles, SPR: the rank of single country articles, CPR: the rank of internationally collaborative articles, FPR: the rank of first author articles, RPR: the rank of corresponding author articles, S%: the percentage of single country articles in each country

the internationally collaborative articles, followed by Germany (1,528) with 23 % and the UK (1,382) with 21 %. Moreover, the USA also had high first author articles (34 %) and corresponding author articles (34 %), and the highest *h*-index of 209. Among the top four countries with more than 2,000 articles, China is the only country whose *h*-index was 67, <100. The phenomenon that China ranked in high position of production but lower in terms of quality assessment was consistent with drinking water field (Fu et al. 2013).

Comparison of the growth trends of the top eight countries is displayed in Fig. 1. Obviously, publications from China grew sharply. This could be due to a series of positive policies motivated the rapid development of research in China. The 11th Five-Year Plan in 2001 and National Medium and Long-term Science and Technology Development Plan in 2006 (in which proteins research acted as one of the four major scientific research projects) was approved by the Chinese government (http://www.most.gov.cn/kjgh). Similar results were also found in the cases of atmospheric simulation research (Li et al. 2009b) and desalination research (Tanaka and Ho 2011).

Institutions' research performances

Publication counts of institutions was a reference for evaluating institutions research performance in a field and had been used in many aspects of research such as lead in drinking water (Hu et al. 2010), and acupuncture (Han and Ho 2011). The contribution of
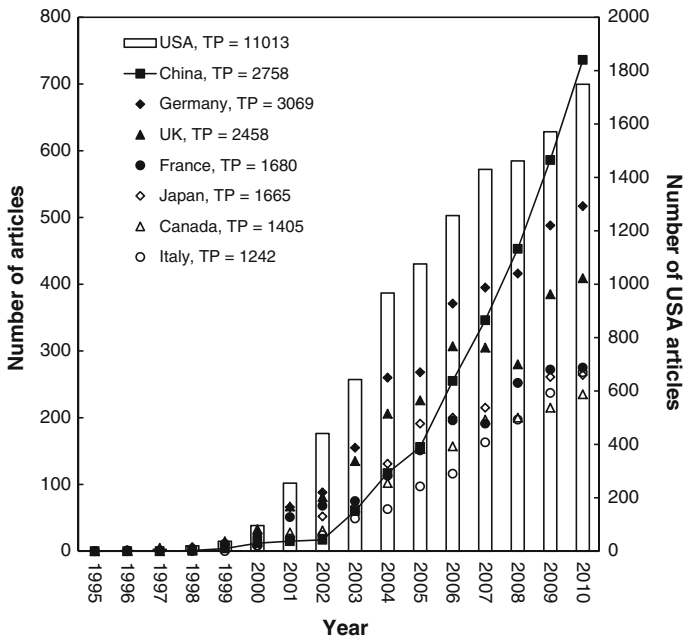
Fig. 1 Comparison the growth trends of the top eight productive countries during 1995–2010

different institutions was estimated by the institution of the affiliation of at least one author of the published articles. Of the 27,774 articles with author addresses, 10,460 (38 %) were single institution articles and 17,314 (62 %) were inter-institutionally collaborative articles. The inter-institutionally collaborative rate was equal to global climate change (Li et al. 2011), but was larger than 53 % of acupuncture research (Han and Ho 2011), 44 % of solid waste research (Fu et al. 2010), and 37 % of desalination research (Tanaka and Ho 2011). The top ten institutions were ranked by the total number of articles, single institution articles, inter-institutionally collaborative articles, first author articles, corresponding author articles, and h-index (Table 4). Among the top ten institutions, six of them were derived from the USA, followed by France with two institutions, and one institution for each of China and South Korea. Leading was the Harvard University of USA with 609 articles and the highest of h-index, which had a large disparity with the others. The Harvard University also published the most single author and collaborative articles, but the second most first author, and corresponding author articles. Although Chinese Academy of Sciences ranked 2nd with respect to publication outputs, its h-index were only 45, even lower than the 10th position University of Texas with h-index of 55. The other two institutions CNRS and INRA also had <50 h-indices. Concerning the trends of top six institutions with more than 300 articles, CNRS started to publish proteomics-related articles since 1996, and remained a mild increase across the decades (Fig. 2). It had high percent of inter-institutionally collaborative articles (rank 4th) and low single-institutionally collaborative articles (rank 52nd). This result demonstrated CNRS had strongly collaborative with other institutions ability in proteomics research. Most dramatic change occurred in USA where the Harvard University started to step into this field in 1999, and became the most published institution in the past 10 years till 2009. Chinese Academy of Sciences began to publish

Table 4 Characteristics of the top ten productive institutions

Institution	TP	TPR (%)	SPR (%)	CPR (%)	FPR (%)	RPR (%)	S%	h-index
Harvard University, USA	609	1 (2.2)	3 (1.1)	1 (2.8)	2 (0.93)	2 (0.91)	28	102
Chinese Academy of Sciences, China	518	2 (1.9)	1 (1.7)	2 (2.0)	1 (1.2)	1 (1.2)	25	45
University of Washington, USA	368	3 (1.3)	9 (0.74)	3 (1.7)	3 (0.78)	3 (0.74)	32	58
Scripps Research Institute, USA	343	4 (1.2)	5 (1.0)	5 (1.4)	5 (0.72)	5 (0.70)	31	75
Pacific Northwest National Laboratory, USA	314	5 (1.1)	2 (1.2)	13 (1.1)	4 (0.75)	4 (0.71)	33	59
CNRS, France	300	6 (1.1)	52 (0.33)	4 (1.5)	22 (0.41)	18 (0.43)	33	45
National Cancer Institute, USA	281	7 (1.0)	19 (0.56)	7 (1.3)	10 (0.56)	10 (0.52)	35	54
University of Michigan, USA	279	8 (1.0)	7 (0.94)	14 (1.0)	6 (0.65)	6 (0.62)	29	55
INRA, France	276	9 (1.0)	20 (0.54)	8 (1.3)	8 (0.58)	9 (0.56)	28	43
University of Toronto, Canada	272	10 (1.0)	32 (0.44)	6 (1.3)	31 (0.38)	31 (0.36)	28	55

TP: number of articles, TPR: the rank of total articles, SPR: the rank of single institution articles, CPR: the rank of inter-institutionally collaborative articles, FPR: the rank of first author articles, RPR: the rank of corresponding author articles, S%: the percentage of single institution articles in each institution

proteomics-related articles in 1999, increased quickly since then, and became the top one after 2009. However, a bias appeared because the Chinese Academy of Sciences had branches in many different cities (Li et al. 2009b).

Research emphases and trends

Distribution of author keywords analysis

Bibliometric methods concerning author keywords had only been found in recent years (Chiu and Ho 2007), and their use in analyzing research trends proved to be useful (Xie et al. 2008; Zhang et al. 2010; Li et al. 2011). Altogether 33,734 were used, 24,949 (74 %) appeared only once and 3,895 (11.5 %) appeared only twice. The large number of once-only used author keywords probably indicated a lack of continuity in research and a wide disparity in research focuses (Chuang et al. 2007). Author keywords that appeared in articles referring on proteomics from 1995 to 2010 were counted and ranked in four 4-year periods (Table 5). Except for "proteomics" and "proteome" which were searching keywords in this study, the three most frequently and stable used keywords were "mass spectrometry" (MS), "two-dimensional gel electrophoresis", and "bioinformatics". All of them were the traditional and essential parts of proteomics research (Boeckmann et al. 2003; Keller et al. 2002), including the separation, identification, and analysis of proteins in a proteome (Yates 1998; Bairoch and Apweiler 2000; Ideker et al. 2001; Gasteiger et al. 2003). With the advance in MS, some new relative technologies came up, such as LC–MS/MS, MS/MS, and MALDI–TOF–MS or MALDI–TOF MS. Their ranks and percentages rose from #543 (0.10 %), #543 (0.10 %), #98 (0.39 %), #144 (0.29 %) in 1999–2002 to
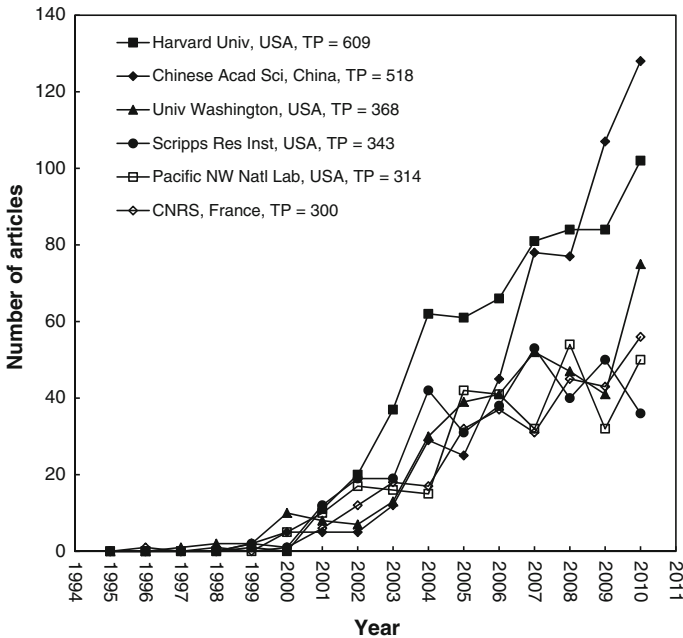
Fig. 2 Comparison the growth trends of the top six productive institutions during 1995–2010

#13 (1.4 %), #99 (0.36 %), #45 (0.43 %), #55 (0.58 %) in 2007–2010 respectively, indicating their increasing importance and popularity in proteomics research (Peng et al. 2003; Elias and Gygi 2007; Gerber et al. 2003; Craig and Beavis 2004).

Meanwhile, the ranks and percentages of articles with "biomarker", "biomarkers", and "oxidative stress" went up, respectively, from #21 (1.3 %), #98 (0.39 %), and #243 (0.20 %) during 1999–2002 to #4 (3.1 %), #7 (2.4 %), and #9 (1.8 %) during 2007–2010. Biomarker was a characteristic that was objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention (Atkinson et al. 2001). Oxidative stress was thought to be involved in the development of many diseases or may exacerbate their symptoms in humans (Proctor 1989; Proctor and Reynolds 1984). Through the proteomics approach, scientists intended to search the biomarkers for early diagnosis, accurate prognosis and prediction of response to treatment in cancer (Simpson and Dorow 2001; Adam et al. 2002), and to analysis oxidative stress in different diseases (Rabilloud et al. 2002; Butterfield 2004; Brennan et al. 2004). On the other hand, the ranks of words such as "two-dimensional polyacrylamide gel electrophoresis", "2D-page", and "2-D page" markedly descended from #1, #38, and #38 during 1995–1998, to #213, #83, and #117 during 2007–2010. The possible explanations for these decreases were that those were general words, and replaced by more specific or definite author keywords, such as two-dimensional gel electrophoresis and 2-DE. The lack of standardization among keywords assigned by authors greatly hampered our analysis since the use of synonymous terms, spelling variations, abbreviations, and more or less specific terms made the exact interpretation of the author's intended meaning difficult (Li et al. 2009b).

Table 5 Top 20 most frequently used author keywords during 1995–2010 and 4-year periods

Words in title	TP	91–10 rank (%)	91–95 rank (%)	96–00 rank (%)	01–05 rank (%)	06–10 rank (%)
Proteomics	4,716	1 (27)	3 (16)	1 (34)	1 (27)	1 (26)
Mass spectrometry	2,116	2 (12)	3 (16)	2 (15)	2 (13)	2 (11)
Proteome	1,133	3 (6.4)	1 (47)	3 (14)	3 (7.1)	3 (5.2)
Two-dimensional gel electrophoresis	764	4 (4.3)	5 (9.4)	4 (11)	4 (6.6)	6 (2.6)
Biomarker	462	5 (2.6)	N/A	21 (1.3)	7 (1.9)	4 (3.1)
Bioinformatics	438	6 (2.5)	10 (4.7)	8 (3.4)	5 (3.5)	8 (1.9)
2-DE	403	7 (2.3)	N/A	N/A	6 (2.2)	5 (2.6)
Biomarkers	372	8 (2.1)	N/A	98 (0.39)	9 (1.9)	7 (2.4)
Two-dimensional electrophoresis	354	9 (2.0)	38 (1.6)	5 (6.4)	8 (1.9)	11 (1.7)
Oxidative stress	299	10 (1.7)	N/A	243 (0.20)	10 (1.8)	9 (1.8)
Apoptosis	274	11 (1.6)	38 (1.6)	21 (1.3)	17 (1.3)	10 (1.7)
Mitochondria	256	12 (1.5)	38 (1.6)	21 (1.3)	15 (1.5)	12 (1.5)
Phosphorylation	231	13 (1.3)	38 (1.6)	35 (0.79)	18 (1.3)	14 (1.4)
Tandem mass spectrometry	221	14 (1.3)	38 (1.6)	14 (2.0)	12 (1.6)	22 (1.0)
Genomics	214	15 (1.2)	8 (6.3)	6 (3.8)	16 (1.4)	26 (0.83)
Breast cancer	207	16 (1.2)	N/A	35 (0.79)	19 (1.2)	17 (1.2)
Protein identification	203	17 (1.2)	5 (9.4)	8 (3.4)	11 (1.6)	43 (0.65)
Proteome analysis	199	18 (1.1)	10 (4.7)	7 (3.5)	13 (1.5)	37 (0.70)
LC–MS/MS	199	18 (1.1)	N/A	543 (0.10)	39 (0.72)	13 (1.4)
Proteomic	199	18 (1.1)	N/A	33 (0.88)	26 (1.0)	15 (1.3)

TP: number of articles, N/A: not available

Contributors to author keywords

To figure out the correlation of author keywords and institutions, three most productive departments of biology, medicine, and chemistry were categorized. Of the 27,774 articles with author addresses, biology departments contributed the most with 18,005 articles (65 %), followed by medicine departments with 12,258 articles (44 %), and chemistry departments with 9,021 articles (32 %). Biologists and medical scientists worked effectively with a great collaboration rate of 30 %, followed by the 26 % of biologists and chemists, and 14 % of medical scientists and chemists. For biologists, approximately a half articles were contributed by the co-operations with medical scientists, and two-fifths articles were collaborated with chemists. To be specific, the relations for the top 25 author keywords and the departments of biology, medicine, and chemistry are identified in Table 6. The descending order of percentages for different departments were biology, medicine, and chemistry, respectively; and for collaboration, the order were accordingly biology and medicine, biology and chemists, and medicine and chemist for 13 author keywords. This performance was consistent with that of total articles. It is noticeable that as for "biomarker", "biomarkers", and "breast cancer", medicine departments contributed more than biology departments, while chemistry departments published more than medicine departments only for "tandem MS", "protein identification", and "proteome analysis". Similarly, biologists have more co-operations with chemists than medical scientists

Table 6 The relations for the 20 most frequently used author keywords and departments of biology, medicine and chemistry

Author keywords	BR (%)	MR (%)	CR (%)	B and MR (%)	B and CR (%)	M and CR (%)	B and M and CR (%)
Proteomics	11 (61)	7 (48)	9 (31)	11 (30)	10 (25)	10 (14)	13 (12)
Mass spectrometry	10 (63)	9 (46)	6 (36)	10 (31)	8 (27)	9 (15)	9 (13)
Proteome	9 (64)	18 (38)	17 (23)	16 (25)	17 (20)	18 (10)	18 (10)
Two-dimensional gel electrophoresis	7 (65)	13 (42)	9 (31)	15 (26)	9 (26)	14 (13)	13 (12)
Biomarker	18 (55)	1 (65)	14 (25)	3 (40)	16 (21)	3 (18)	3 (16)
Bioinformatics	1 (73)	6 (51)	13 (26)	1 (43)	13 (23)	10 (14)	9 (13)
2-DE	3 (68)	10 (44)	14 (25)	7 (32)	13 (23)	17 (12)	17 (11)
Biomarkers	19 (54)	2 (61)	17 (23)	4 (38)	18 (19)	10 (14)	9 (13)
Two-dimensional electrophoresis	11 (61)	15 (40)	16 (24)	16 (25)	11 (24)	14 (13)	13 (12)
Oxidative stress	11 (61)	7 (48)	3 (38)	6 (34)	7 (28)	1 (20)	2 (17)
Apoptosis	3 (68)	3 (60)	12 (27)	2 (41)	11 (24)	5 (17)	4 (15)
Mitochondria	7 (65)	13 (42)	5 (37)	12 (29)	4 (30)	7 (16)	4 (15)
Phosphorylation	5 (66)	5 (52)	8 (35)	7 (32)	4 (30)	5 (17)	4 (15)
Tandem mass spectrometry	17 (57)	15 (40)	1 (48)	18 (24)	1 (33)	3 (18)	7 (14)
Genomics	20 (47)	15 (40)	20 (16)	20 (18)	20 (13)	20 (6.1)	20 (5.6)
Breast cancer	16 (58)	4 (59)	11 (28)	4 (38)	15 (22)	7 (16)	7 (14)
Protein identification	11 (61)	19 (34)	2 (41)	14 (27)	4 (30)	10 (14)	9 (13)
Proteome analysis	5 (66)	20 (31)	6 (36)	19 (21)	2 (32)	14 (13)	13 (12)
LC–MS/MS	2 (69)	10 (44)	3 (38)	7 (32)	2 (32)	1 (20)	1 (18)
Proteomic	15 (59)	10 (44)	17 (23)	12 (29)	18 (19)	19 (8.0)	19 (6.5)

BR: the rank of articles published by biology departments, MR: the rank of articles published by medicine departments, CR: the rank of articles published by chemistry departments, B and MR: the rank of articles published by biology and medicine departments, B and CR: the rank of articles published by biology and chemistry departments, M and CR: the rank of articles published by biology and chemistry departments, B and M and CR: the rank of articles published by biology, medicine, and chemistry departments

for five words including "mitochondria", "tandem MS", "protein identification", "proteome analysis", and "LC–MS/MS". In particular, chemists were likely to work together with biologist and medical scientist, instead of working independently in proteomics research.

Hot issues

Based on the important clues from author keywords analysis, the research emphases were summed by some synonymic single words and congeneric phrases to analyze the historical development of the science (Mao et al. 2010). Research trends in proteomics were separated into two categories research methods including two dimensional electrophoresis (2DE) and MS, and affected items including cancer and plants. The topic 2DE contained the words "2 (-) dimensional electrophoresis", "2 (-) DE" or "2d (-) page", "2 (-) D dige", "2 (-) d electrophoresis", "two dimensional gel electrophoresis", "2 (-) dimensional gel electrophoresis", "two dimensional polyacrylamide gel electrophoresis", "2 (-)

dimensional polyacrylamide gel electrophoresis", and "difference gel electrophoresis". The topic MS contained the words "MS" and "mass spectroscopy". The topic cancer contained the words "cancer", "carcinoma", "tumor", and "metastasis", and the topic plants contained the words "rice", *Oryza sativa*", "arabidopsis", "wheat", "soybean", "maize", "tomato", *Medicago truncatula*", *Zea mays*", "grape" and "sod". 2DE and MS were two most important and standard technologies in proteomics research (Gygi et al. 2000; Imai and Mische 1999). 2DE was a powerful separation technique, which allows simultaneous resolution of thousands of proteins (Görg et al. 2004), and MS had increasingly become the method for analysis of complex protein samples (Aebersold and Mann 2003; Mann et al. 2001). At the end of 2010, both of them had distinctly high incidence, being mentioned in 6,728 and 13,786 articles respectively. "Cancer proteomics" was the analysis of molecular pathogenesis of cancer by analyzing global protein expression in tumors, tumor cells or extracellular fluids from cancer patients (Blackstock and Weir 1999), and it had been widely applied to cancer research (Kolch and Pitt 2010). There was an increasing number of reports on the clinical application of proteomics research for risk assessment, diagnosis, prognosis, and management of cancer (Zhang and Chan 2005). For plants, proteomics was becoming increasingly important for the study of many different aspects of plant functions, revealing molecular mechanisms underlying plant growth, development, and interactions with the environment (Chen and Harmon 2006). It will ultimately enable predictable engineering of plant processes that were important to crop yield, nutrition, and defense (Nelson et al. 2006; Macherel et al. 2007; Jung et al. 2006; Lee et al. 2007).

Figure 3 showed the research trends for four topics of "2DE", "MS", "cancer", and "plants" from 1995 to 2010. The number of articles related MS increased rapidly from 2002, while one related 2DE increased slower or decreased after 2006. This was because that 2DE was an old fashioned technology with many limitations (Cellulaire 2002), and a great deal of progress on MS had been made in the last 10 years (Wolters et al. 2001; Clauser et al. 1999). While 2DE was introduced in the mid-70s (Macgillivray and Wood 1974; Klose 1975; O'Farrell 1975), its real expansion as a useful technique had to wait for the development of microanalytical techniques able to identify proteins at the amounts available from 2-D gels. MS had greatly increased the power of this microcharacterization step (Gygi et al. 2000). With the increased performance and versatility of the proteomics instrumentation, new protein analytical strategies had emerged in which MS was the central element, such as matrix-assisted laser desorption/ionization time-of-light MS (MALDI–TOF–MS) and electrospray ionization-quadrupole ion trap MS (ESI–Q–IT–MS) (Domon and Aebersold 2006; Shevchenko et al. 2001; Kinter and Sherman 2000). Now, MS was joined by antibody and protein–protein interaction arrays (Rual et al. 2005; Wolf-Yadlin et al. 2009), fluorescence-and flow cytometry-based detection of proteins, post-translational modifications (PTMs) (Schulz et al. 2007), and optical spectroscopic methods of proteome analysis (Fournier et al. 2009; Faley et al. 2009). Therefore, it was expected that MS would grow constantly in the next decades. Referring to the items about cancer and plants in proteomics, more attention was paid to the research on "cancer". The number of articles related to "cancer" exceeded that of "plants" after 1998, being mentioned in 6,212 articles at a rate of over 100 articles in 2003, 2004, 2005, 2008, 2009, and 2010. For most cancers, survival rates depended on the early detection of the disease (Seibert et al. 2005). Proteomics technologies had brought the hope of discovering novel early cancer-specific biomarkers in complex biological samples (Gygil et al. 1999). Moreover, it helped us gain functional understanding of the pathways that were deranged in cancer (Kolch and Pitt 2010). Global-expression proteomics (Celis et al. 1999; Celis et al. 2000) and cell-
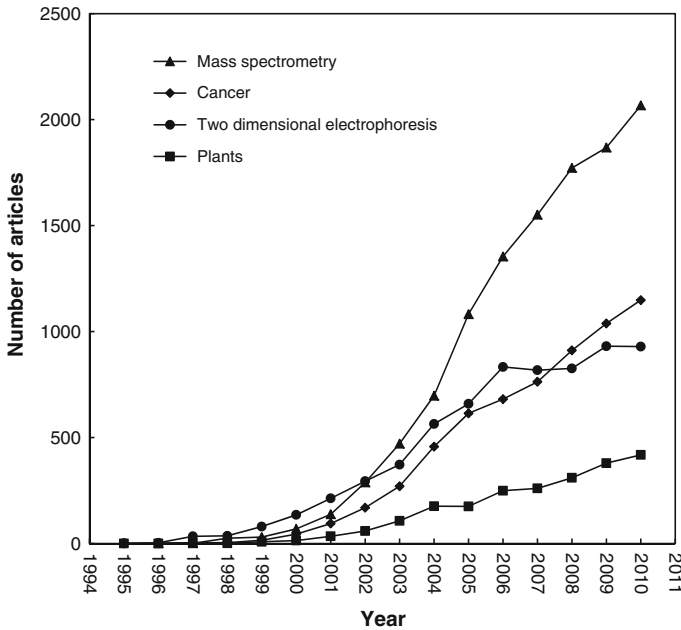
Fig. 3 Growth trends of hotspot-related articles in the recent 16 years

mapping proteomics (Lewis et al. 2000; Verhagen et al. 2000; Zhang et al. 1999) will contribute to significant advances in understanding cancer. In plants proteomics, after the first plant large-scale proteomic work (Kamo et al. 1995), the full potential of proteomics was far from being exploited, especially compared to humans. Even so, proteomics was providing new information, validating, complementing or correcting that information already provided by different approaches, as a consequence, giving us a deeper knowledge of plant biology (Jorrín et al. 2007).

Conclusion

This study on proteomics dealing with *SCI-Expanded* obtained some significant points on the research performance throughout the period from 1995 to 2010. *Proteomics* and *Journal of Proteome Research* were the most productive journals. The categories of biochemical research methods, and biochemistry and molecular biology hold about one-third of proteomics research, respectively. The USA took the lead in total publications, while China had the quickest increasing based on its national policy. The most productive institution was Harvard University, but one China's institution, Chinese Academy of Science, surpassed it in 2010. However, the lower *h*-indices of China and Chinese Academy of Science indicated that the quality of research needs to be improved. Furthermore, the innovatively analyzed distribution of author keywords provided important clues for research emphases. The trend analysis of hot issues based on title words, author keywords, abstract words, and *KeyWords Plus* indicated that 2-DE and MS were the two major technology of proteomics research, and proteomics research related to cancer will continue being the emphases in the foreseeable future. Biologist contributed the most to proteomics research, and had significant co-operation with medical scientists and chemists.

References

Adam, B. L., Qu, Y., Davis, J. W., Ward, M. D., Clements, M. A., Cazares, L. H., et al. (2002). Serum protein fingerprinting coupled with a pattern-matching algorithm distinguishes prostate cancer from benign prostate hyperplasia and healthy men. *Cancer Research, 62*(13), 3609–3614.

Aebersold, R., & Mann, M. (2003). Mass spectrometry-based proteomics. *Nature, 422*(6928), 198–207.

Alfaraz, P. H., & Calvino, A. M. (2004). Bibliometric study on food science and technology: Scientific production in Iberian–American countries (1991–2000). *Scientometrics, 1*, 89–102.

Allen, T. M., & Cullis, P. R. (2004). Drug delivery systems: Entering the mainstream. *Science, 303*(5665), 1818–1822.

Alm, R. A., Ling, L. S. L., Moir, D. T., King, B. L., Brown, E. D., Doig, P. C., et al. (1999). Genomic-sequence comparison of two unrelated isolates of the human gastric pathogen *Helicobacter pylori*. *Nature, 397*(6715), 176–180.

Anderson, N. (1998). Proteome and proteomics: New technologies, new concepts, and new words. *Electrophoresis, 19*(11), 1853–1861.

Atkinson, A. J., Colburn, W. A., DeGruttola, V. G., DeMets, D. L., Downing, G. J., Hoth, D. F., et al. (2001). Biomarkers and surrogate endpoints: Preferred definitions and conceptual framework. *Clinical Pharmacology and Therapeutics, 69*(3), 89–95.

Bairoch, A., & Apweiler, R. (2000). The SWISS–PROT protein sequence database and its supplement TrEMBL in 2000. *Nucleic Acids Research, 28*(1), 45–48.

Bajwa, R. S., Yaldram, K., & Rafique, S. (2013). A scientometric assessment of research output in nanoscience and nanotechnology: Pakistan perspective. *Scientometrics, 94*(1), 333–342.

Bayer, A. E., & Folger, J. (1966). Some correlates of a citation measure of productivity in science. *Sociology of Education, 39*(4), 381–390.

Blackstock, W. P., & Weir, M. P. (1999). Proteomics: Quantitative and physical mapping of cellular proteins. *Trends in Biotechnology, 17*(3), 121–127.

Boeckmann, B., Bairoch, A., Apweiler, R., Blatter, M. C., Estreicher, A., Gasteiger, E., et al. (2003). The SWISS–PROT protein knowledgebase and its supplement TrEMBL in 2003. *Nucleic Acids Research, 31*(1), 365–370.

Bradford, S. C. (1934). Sources of information on specific subjects. *British Journal of Engineering, 137*(3550), 85–86.

Braun, T., Schubert, A. P., & Kostoff, R. N. (2000). Growth and trends of fullerene research as reflected in its journal literature. *Chemical Reviews, 100*(1), 23–38.

Brennan, J. P., Wait, R., Begum, S., Bell, J. R., Dunn, M. J., & Eaton, P. (2004). Detection and mapping of widespread intermolecular protein disulfide formation during cardiac oxidative stress using proteomics with diagonal electrophoresis. *Journal of Biological Chemistry, 279*(40), 41352–41360.

Butterfield, D. A. (2004). Proteomics: A new approach to investigate oxidative stress in Alzheimer's disease brain. *Brain Research, 1000*(1), 1–7.

Celis, J. E., Ostergaard, M., Rasmussen, H. H., Gromov, P., Gromova, I., Varmark, H., et al. (1999). A comprehensive protein resource for the study of bladder cancer: http://biobase.dk/cgi-bin/celis. *Electrophoresis*, *20*(2), 300–309.

Celis, J. E., Wolf, H., & Østergaard, M. (2000). Bladder squamous cell carcinoma biomarkers derived from proteomics. *Electrophoresis, 21*(11), 2115–2121.

Cellulaire, B. (2002). Two-dimensional gel electrophoresis in proteomics: Old, old fashioned, but it still climbs up the mountains. *Proteomics, 2*, 3–10.

Chen, S., & Harmon, A. C. (2006). Advances in plant proteomics. *Proteomics, 6*(20), 5504–5516.

Chiu, W. T., & Ho, Y. S. (2005). Bibliometric analysis of homeopathy research during the period of 1991 to 2003. *Scientometrics, 63*(1), 3–23.

Chiu, W. T., & Ho, Y. S. (2007). Bibliometric analysis of tsunami research. *Scientometrics, 73*(1), 3–17.

Chuang, K. Y., Huang, Y. L., & Ho, Y. S. (2007). A bibliometric and citation analysis of stroke-related research in Taiwan. *Scientometrics, 72*(2), 201–212.

Chuang, K. Y., Wang, M. H., & Ho, Y. S. (2011). High-impact papers presented in the subject category of water resources in the essential science indicators database of the Institute for Scientific Information. *Scientometrics, 87*(3), 551–562.

Clauser, K. R., Baker, P., & Burlingame, A. L. (1999). Role of accurate mass measurement (±10 ppm) in protein identification strategies employing MS or MS/MS and database searching. *Analytical Chemistry, 71*(14), 2871–2882.

Coats, A. J. S. (2009). Ethical authorship and publishing. *International Journal of Cardiology, 131*(2), 149–150.

Craig, R., & Beavis, R. C. (2004). TANDEM: Matching proteins with tandem mass spectra. *Bioinformatics, 20*(9), 1466–1467.

Domon, B., & Aebersold, R. (2006). Mass spectrometry and protein analysis. *Science, 312*(5771), 212–217.

Elias, J. E., & Gygi, S. P. (2007). Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nature Methods, 4*(3), 207–214.

Faley, S. L., Copland, M., Wlodkowic, D., Kolch, W., Seale, K. T., Wikswo, J. P., et al. (2009). Microfluidic single cell arrays to interrogate signalling dynamics of individual, patient-derived hematopoietic stem cells. *Lab on a Chip, 9*(18), 2659–2664.

Fournier, F., Guo, R., Gardner, E. M., Donaldson, P. M., Loeffeld, C., Gould, I. R., et al. (2009). Biological and biomedical applications of two-dimensional vibrational spectroscopy: proteomics, imaging, and structural analysis. *Accounts of Chemical Research, 42*(9), 1322–1331.

Fu, H. Z., Ho, Y. S., Sui, Y. M., & Li, Z. S. (2010). A bibliometric analysis of solid waste research during the period 1993–2008. *Waste Management, 30*(12), 2410–2417.

Fu, H. Z., Wang, M. H., & Ho, Y. S. (2013). Mapping of drinking water research: A bibliometric analysis of research output during 1992–2011. *Science of the Total Environment, 443*, 757–765.

Garfield, E. (1990). KeyWords Plus^{TM}—ISIS breakthrough retrieval method. 1. Expanding your searching power on current-contents on diskette. *Current Contents, 32*, 5–9.

Garrels, J., McLaughlin, C., Warner, J., Futcher, B., Latter, G., Kobayashi, R., et al. (1997). Proteome studies of *Saccharomyces cerevisiae*: Identification and characterization of abundant proteins. *Electrophoresis, 18*(8), 1347–1360.

Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R. D., & Bairoch, A. (2003). ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Research, 31*(13), 3784–3788.

Gavin, A. C., Bösche, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., et al. (2002). Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature, 415*(6868), 141–147.

Gerber, S. A., Rush, J., Stemman, O., Kirschner, M. W., & Gygi, S. P. (2003). Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proceedings of the National Academy of Sciences, 100*(12), 6940–6945.

Görg, A., Weiss, W., & Dunn, M. J. (2004). Current two-dimensional electrophoresis technology for proteomics. *Proteomics, 4*(12), 3665–3685.

Gygi, S. P., Corthals, G. L., Zhang, Y., Rochon, Y., & Aebersold, R. (2000). Evaluation of two-dimensional gel electrophoresis-based proteome analysis technology. *Proceedings of the National Academy of Sciences, 97*(17), 9390–9395.

Gygil, S. P., Rist, B., Gerber, S. A., Turecek, F., Gelb, M. H., & Aebersold, R. (1999). Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nature Biotechnology, 17*(10), 994–999.

Han, J. S., & Ho, Y. S. (2011). Global trends and performances of acupuncture research. *Neuroscience and Biobehavioral Reviews, 35*(3), 680–687.

Hanash, S. (2003). Disease proteomics. *Nature, 422*(6928), 226–232.

Harris, L., Fritsche, H., Mennel, R., Norton, L., Ravdin, P., Taube, S., et al. (2007). American Society of Clinical Oncology 2007 update of recommendations for the use of tumor markers in breast cancer. *Journal of Clinical Oncology, 25*(33), 5287–5312.

Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences of the United States of America, 102*(46), 16569–16572.

Ho, Y. S. (2008). Bibliometric analysis of biosorption technology in water treatment research from 1991 to 2004. *International Journal of Environment and Pollution, 34*(1–4), 1–13.

Ho, Y. S., Satoh, H., & Lin, S. Y. (2010). Japanese lung cancer research trends and performances in Science Citation Index. *Internal Medicine, 49*(20), 2219–2228.

Hood, L., Heath, J. R., Phelps, M. E., & Lin, B. (2004). Systems biology and new technologies enable predictive and preventative medicine. *Science, 306*(5696), 640–643.

Hu, J., Ma, Y. W., Zhang, L., Gan, F. X., & Ho, Y. S. (2010). A historical review and bibliometric analysis of research on lead in drinking water field from 1991 to 2007. *Science of the Total Environment, 408*(7), 1738–1744.

Ideker, T., Thorsson, V., Ranish, J. A., Christmas, R., Buhler, J., Eng, J. K., et al. (2001). Integrated genomic and proteomic analyses of a systematically perturbed metabolic network. *Science, 292*(5518), 929–934.

Imai, B. S., & Mische, S. M. (1999). Mass spectrometric identification of proteins from silver-stained polyacrylamide gel: A method for the removal of silver ions to enhance sensitivity. *Electrophoresis, 20*, 601–605.

James, P. (1997). Protein identification in the post-genome era: The rapid rise of proteomics. *Quarterly Reviews of Biophysics, 30*(4), 279–331.

Jorrín, J. V., Maldonado, A. M., & Castillejo, M. A. (2007). Plant proteome analysis: A 2006 update. *Proteomics, 7*(16), 2947–2962.

Jung, Y. H., Rakwal, R., Agrawal, G. K., Shibato, J., Kim, J. A., Lee, M. O., et al. (2006). Differential expression of defense/stress-related marker proteins in leaves of a unique rice blast lesion mimic mutant (BLM). *Journal of Proteome Research, 5*(10), 2586–2598.

Kamo, M., Kawakami, T., Miyatake, N., & Tsugita, A. (1995). Separation and characterization of *Arabidopsis thaliana* proteins by two-dimensional gel electrophoresis. *Electrophoresis, 16*(3), 423–430.

Keller, A., Nesvizhskii, A. I., Kolker, E., & Aebersold, R. (2002). Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Analytical Chemistry, 74*(20), 5383–5392.

Kinter, M., & Sherman, N. E. (2000). *Protein sequencing and identification using tandem mass spectrometry* (Vol. 2). New York: Wiley.

Klose, J. (1975). Protein mapping by combined isoelectric focusing and electrophoresis of mouse tissues. *Human Genetics, 26*(3), 231–243.

Kolch, W., & Pitt, A. (2010). Functional proteomics to dissect tyrosine kinase signalling pathways in cancer. *Nature Reviews Cancer, 10*(9), 618–629.

Lee, D. G., Ahsan, N., Lee, S. H., Kang, K. Y., Bahk, J. D., Lee, I. J., et al. (2007). A proteomic approach in analyzing heat-responsive proteins in rice leaves. *Proteomics, 7*(18), 3369–3383.

Lewis, T. S., Hunt, J. B., Aveline, L. D., Jonscher, K. R., Louie, D. F., Yeh, J. M., et al. (2000). Identification of novel MAP kinase pathway signaling targets by functional proteomics and mass spectrometry. *Molecular Cell, 6*(6), 1343–1354.

Li, L. L., Ding, G. H., Feng, N., Wang, M. H., & Ho, Y. S. (2009a). Global stem cell research trend: Bibliometric analysis as a tool for mapping of trends from 1991 to 2006. *Scientometrics, 80*(1), 39–58.

Li, J. F., Wang, M. H., & Ho, Y. S. (2011). Trends in research on global climate change: A Science Citation Index Expanded-based analysis. *Global and Planetary Change, 77*(1–2), 13–20.

Li, J., Zhang, Z., Rosenzweig, J., Wang, Y. Y., & Chan, D. W. (2002). Proteomics and bioinformatics approaches for identification of serum biomarkers to detect breast cancer. *Clinical Chemistry, 48*(8), 1296–1304.

Li, J. F., Zhang, Y. H., Wang, X. S., & Ho, Y. S. (2009b). Bibliometric analysis of atmospheric simulation trends in meteorology and atmospheric science journals. *Croatica Chemica Acta, 82*(3), 695–705.

Link, A., Hays, L., Carmack, E., & Yates, J., I. I. I. (1997). Identifying the major proteome components of *Haemophilus influenzae* type-strain NCTC 8143. *Electrophoresis, 18*(8), 1314–1334.

Macgillivray, A. J., & Wood, D. R. (1974). The heterogeneity of mouse-chromatin nonhistone proteins as evidenced by two-dimensional polyacrylamide-gel electrophoresis and ion-exchange chromatography. *European Journal of Biochemistry, 41*(1), 181–190.

Macherel, D., Benamar, A., Avelange-Macherel, M. H., & Tolleter, D. (2007). Function and stress tolerance of seed mitochondria. *Physiologia Plantarum, 129*(1), 233–241.

Mann, M., Hendrickson, R. C., & Pandey, A. (2001). Analysis of proteins and proteomes by mass spectrometry. *Annual Review of Biochemistry, 70*(1), 437–473.

Mao, N., Wang, M. H., & Ho, Y. S. (2010). A bibliometric study of the trend in articles related to risk assessment published in Science Citation Index. *Human and Ecological Risk Assessment, 16*(4), 801–824.

Moed, H. F., Burger, W. J. M., Frankfort, J. G., & Vanraan, A. F. J. (1985). The use of bibliometric data for the measurement of university research performance. *Research Policy, 14*(3), 131–149.

Muellner, S., Neumann, T., & Lottspeich, F. (1998). Proteomics—A new way for drug target discovery. *Arzneimittel-Forschung, 48*(1), 93–95.

Nelson, T., Tausta, S. L., Gandotra, N., & Liu, T. (2006). Laser microdissection of plant tissue: What you see is what you get. *Annual Review of Plant Biology, 57*, 181–201.

O'Farrell, P. H. (1975). High resolution two-dimensional electrophoresis of proteins. *Journal of Biological Chemistry, 250*(10), 4007–4021.

Ong, S. E., & Mann, M. (2005). Mass spectrometry-based proteomics turns quantitative. *Nature Chemical Biology, 1*(5), 252–262.

Pandey, A., & Mann, M. (2000). Proteomics to study genes and genomes. *Nature, 405*(6788), 837–846.

Peng, J., Elias, J. E., Thoreen, C. C., Licklider, L. J., & Gygi, S. P. (2003). Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC–MS/MS) for large-scale protein analysis: The yeast proteome. *Journal of Proteome Research, 2*(1), 43–50.

Persidis, A. (1998). Proteomics—An ambitious drug development platform attempts to link gene sequence to expressed phenotype under various physiological states. *Nature Biotechnology, 16*(4), 393–394.

Pritchard, A. (1969). Statistical bibliography or bibliometrics. *Journal of Documentation, 25*(4), 348–349.

Proctor, P. H. (1989). Free radicals and human disease. *CRC Handbook of Free Radicals and Antioxidants, 1*, 209–221.

Proctor, P. H., & Reynolds, E. (1984). Free radicals and disease in man. *Physiological Chemistry and Physics and Medical NMR, 16*(3), 175–195.

Qi, S. Y., Moir, A., & O'Connor, C. D. (1996). Proteome of *Salmonella typhimurium* SL1344: Identification of novel abundant cell envelope proteins and assignment to a two-dimensional reference map. *Journal of Bacteriology, 178*(16), 5032–5038.

Rabilloud, T., Heller, M., Gasnier, F., Luche, S., Rey, C., Aebersold, R., et al. (2002). Proteomics analysis of cellular response to oxidative stress. *Journal of Biological Chemistry, 277*(22), 19396–19401.

Rifai, N., Gillette, M. A., & Carr, S. A. (2006). Protein biomarker discovery and validation: The long and uncertain path to clinical utility. *Nature Biotechnology, 24*(8), 971–983.

Rual, J. F., Venkatesan, K., Hao, T., Hirozane-Kishikawa, T., Dricot, A., Li, N., et al. (2005). Towards a proteome-scale map of the human protein–protein interaction network. *Nature, 437*(7062), 1173–1178.

Schubert, A., Glänzel, W., & Braun, T. (1989). Scientometric datafiles: A comprehensive set of indicators on 2649 journals and 96 countries in all major science fields and subfields 1981–1985. *Scientometrics, 16*, 3–478.

Schulz, K. R., Danna, E. A., Krutzik, P. O., & Nolan, G. P. (2007). Single-cell phospho-protein analysis by flow cytometry. *Current Protocols in Immunology, 8*(17), 11–18.

Seibert, V., Ebert, M. P. A., & Buschmann, T. (2005). Advances in clinical cancer proteomics: SELDI-ToF-mass spectrometry and biomarker discovery. *Briefings in Functional Genomics & Proteomics, 4*(1), 16–26.

Shevchenko, A., Jensen, O. N., Podtelejnikov, A. V., Sagliocco, F., Wilm, M., Vorm, O., et al. (1996). Linking genome and proteome by mass spectrometry: Large-scale identification of yeast proteins from two dimensional gels. *Proceedings of the National Academy of Sciences, 93*(25), 14440–14445.

Shevchenko, A., Sunyaev, S., Loboda, A., Bork, P., & Ens, W. (2001). Charting the proteomes of organisms with unsequenced genomes by MALDI-quadrupole time-of-flight mass spectrometry and BLAST homology searching. *Analytical Chemistry, 73*(9), 1917–1926.

Simpson, R. J., & Dorow, D. S. (2001). Cancer proteomics: From signaling networks to tumor markers. *Trends in Biotechnology, 19*, 40–48.

Sun, J. S., Wang, M. H., & Ho, Y. S. (2012). A historical review and bibliometric analysis of research on estuary pollution. *Marine Pollution Bulletin, 64*(1), 13–21.

Tanaka, H., & Ho, Y. S. (2011). Global trends and performances of desalination research. *Desalination and Water Treatment, 25*(1–3), 1–12.

Tyers, M., & Mann, M. (2003). From genomics to proteomics. *Nature, 422*, 193–197.

Verhagen, A. M., Ekert, P. G., Pakusch, M., Silke, J., Connolly, L. M., Reid, G. E., et al. (2000). Identification of DIABLO, a mammalian protein that promotes apoptosis by binding to and antagonizing IAP proteins. *Cell, 102*(1), 43–53.

Wang, M. H., & Ho, Y. S. (2011). Research articles and publication trends in environmental sciences from 1998 to 2009. *Archives of Environmental Science, 5*, 1–10.

Washburn, M. P., Wolters, D., & Yates, J. R., I. I. I. (2001). Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nature Biotechnology, 19*(3), 242–247.

Wen, D., Yu, T. C., & Ho, Y. S. (2009). Bibliometric tools applied to analytical articles: The example of gene transfer-related research. *OCLC Systems & Services, 25*(3), 186–199.

Wilkins, M. R., Pasquali, C., Appel, R. D., Ou, K., Golaz, O., Sanchez, J. C., et al. (1996). From proteins to proteomes: Large scale protein identification by two-dimensional electrophoresis and arnino acid analysis. *Nature Biotechnology, 14*(1), 61–65.

Wolf-Yadlin, A., Sevecka, M., & MacBeath, G. (2009). Dissecting protein function and signaling using protein microarrays. *Current Opinion in Chemical Biology, 13*(4), 398–405.

Wolters, D. A., Washburn, M. P., & Yates, J. R., I. I. I. (2001). An automated multidimensional protein identification technology for shotgun proteomics. *Analytical Chemistry, 73*(23), 5683–5690.

Xie, S. D., Zhang, J., & Ho, Y. S. (2008). Assessment of world aerosol research trends by bibliometric analysis. *Scientometrics, 77*(1), 113–130.

Yan, J., Tonella, L., Sanchez, J., Wilkins, M., Packer, N., Gooley, A., et al. (1997). The *Dictyostelium discoideum* proteome—The SWISS-2DPAGE database of the multicellular aggregate (slug). *Electrophoresis, 18*(3–4), 491–497.

Yates, J. R. (1998). Mass spectrometry and the age of the proteome. *Journal of Mass Spectrometry, 33*(1), 1–19.

Yu, J. J., Wang, M. H., Xu, M., & Ho, Y. S. (2012). A bibliometric analysis of research papers published on photosynthesis: 1992–2009. *Photosynthetica, 51*(1), 5–14.

Zhang, Z., Bast, R. C., Yu, Y., Li, J., Sokoll, L. J., Rai, A. J., et al. (2004). Three biomarkers identified from serum proteomic analysis for the detection of early stage ovarian cancer. *Cancer Research, 64*(16), 5882–5890.

Zhang, Z., & Chan, D. W. (2005). Cancer proteomics: In pursuit of "true" biomarker discovery. *Cancer Epidemiology, Biomarkers and Prevention, 14*(10), 2283–2286.

Zhang, J. G., Farley, A., Nicholson, S. E., Willson, T. A., Zugaro, L. M., Simpson, R. J., et al. (1999). The conserved SOCS box motif in suppressors of cytokine signaling binds to elongins B and C and may couple bound proteins to proteasomal degradation. *Proceedings of the National Academy of Sciences, 96*(5), 2071–2076.

Zhang, G. F., Xie, S. D., & Ho, Y. S. (2010). A bibliometric analysis of world volatile organic compounds research trends. *Scientometrics, 83*(2), 477–492.