

체질임상정보 및 바이오정보를 활용한 연구동향분석에 대한 연구

한국한의학연구원 한의학정보화사업단
김철 · 송미영 · 엄동명*

A study of Tendency Analysis to Research Using Clinical and Biological Information

Korea Medicine Information Division, Korea Institute of Oriental Medicine
Kim, Chul · Song, Mi-Young · Eom, Dong-Myung

The fusion research subjects of Oriental medicine and Bio-Technology are actively advanced. Development of information analysis algorithm for constitution diagnosis using clinical and biological information is started recently. In addition, the research for providing personal oriented treatment information with Diagnosis result is to in the process of advancing. This paper considers the inside and outside of the country technical development trend of research for using clinical and biological information by analyzing patents and papers.

It executed information analysis about changes to number of research papers and patents, present state and star higher officer of research facility from the dissertation which it sees. It is known that our country research result is slight so far in quantity and quality as result of analysis. But hereafter it contains a many developmental possibility. Also it reflects the appearance and a growth of new field like bioinformatics from constitution and bio information application field. The combination of bio information similar to constitution and heredity and computing technology to control those information efficiently comes to think with core technique.

Key Words : Information Analysis, Paper Analysis, Patent Analysis, Bioinformatics

I. 서론

* 교신저자 : 엄동명, 한국한의학연구원 한의학정보화사업단,
042-868-9450, haksan@kiom.re.kr

* 이 논문은 기초기술연구회 '06협동연구사업 '체질임상정보와 바이오정보를 활용한 체질특성 정보분석 알고리즘 개발'(C-연구-2006-10-KIOM)과제의 지원을 받아 연구되었음.

요즈음 한의학과 BT가 융합된 연구과제들이 활발히 진행되고 있다. 한의학 고전적인 임상 데이터와 바이오정보인 SNP(Single Nucleotide Polymorphism : 단일

염기다형성 현상) 또는 단백질 profile data를 이용하여 체질을 진단하고, 진단결과를 활용하여 맞춤형 치료 정보를 제공하기 위한 연구가 진행 중에 있으며 한의학 관점에서 바이오인포매틱스 Tool 개발을 최종적인 목표로 하는 연구도 추진되고 있다. 바이오인포매틱스(Bioinformatics)는 생명공학기술(Biotechnology)과 정보기술(Information Technology)이 융합되어 생명현상 연구에서 나오는 다양한 생물정보를 다루는 분야로서 기존 바이오인포매틱스가 정의하는 영역에 한의학을 접목시키는 관점의 연구가 시도되고 있고 그러한 연구동향에 대한 정보분석이 필요한 시점이다.

본 논문에서는 논문 및 특허 자료를 활용한 정보 분석을 통하여 생물정보학적 관점에서 임상정보, SNP 정보, 단백질 정보를 바탕으로 체질을 과학적으로 판별하며 재현 가능한 정보분석 도구를 개발하는 과제 및 이와 유사한 한의학적 임상데이터와 바이오데이터간의 융합연구에 대한 국내의 기술개발 동향을 파악하였다. 정보분석을 실시하는 목적은 현재 관련된 연구의 성숙도 정도를 알아보고, 유사 연구를 수행하는 경쟁자의 동향을 파악하며, 기술추이 분석 및 미래예측을 하여 연구개발 방향을 재설정하고자 하는데 있다. 또한 이러한 정보분석 과정은 올해부터 많은 기관에서 신규과제에 특허 및 논문 분석을 의무화할 정도로 중요하다고 판단된다.

II. 연구방법

분석의 주제는 체질임상정보와 바이오정보를 활용한 연구동향이며, 체질임상정보 및 바이오정보를 활용한 연구동향을 알기 위해 관련 키워드로 검색하여 얻은 학술논문정보 및 특허정보를 토대로 메가트렌드분석, 국가 위상분석, 연구주체(연구자 및 연구기관) 분석 등을 실시하였다.

분석인자	논문	논문발표연도, 저자, 논문저자의 국적 및 소속기관, 기술분류, 논문키워드, 논문 피인용횟수, 논문제목과 초록정보, 저널명 등
	특허	특허출원연도 및 출원국가, 기본특허(Basic Patent)의 연도 및 국가, 우선권주장국 및 우선권주장연도, 출원인, 발명자, 발명자의 국적, 기술분류(Derwent Classification), 국제특허분류(International Patent Classification), 특허 피인용수, 특허제목(발명의 명칭)과 초록정보 등
분석내용		◎ 체질임상정보 및 바이오정보 활용분야 국내외 연구동향 분석 -시계열 추이분석, 국가분석, 소속기관(연구기관, 출원인) 분석 -저자(연구자, 발명자) 분석, 연구분야 분석 ◎ 네트워크 분석 -국제공동연구 네트워크, 연구분야간 네트워크 등

【표 1】 체질임상정보 및 바이오정보 활용연구에 대한 정보분석인자 및 분석내용

분석대상기간은 논문의 경우 1987~2006년의 20년간으로 한정하였으며, 특허는 1977~2006년의 30년간으로 하였다. 세부적인 분석내용은 표 1과 같다.

체질임상정보 및 바이오정보를 활용한 연구성과를 수집하기 위해 우선 전문가와 협의하여 관련정보를 최대한 수집할 수 있는 검색어를 선정하였으며, 이를 활용하여 1987~2006년에 발표된 관련 논문정보를 Web of Science 데이터베이스로부터, 그리고 1977~2006년에 등록된 특허정보를 DWPI(Derwent World Patent Index) 데이터베이스로부터 검색수집하여 분석대상으로 하였다.

자료 검색 및 수집에 사용된 검색어는 표 2와 같다. 검색결과로 얻은 4,633편의 논문 중 데이터베이스의 주제 카테고리(Subject category)에 적합하지 않은 논문을 검색노이즈로 판단, 제거하여 최종적으로는 4,404편의 논문을 정보분석에 이용하였다. 검색노이즈는 DB에 저장되어 있는 분류체계를 준용하여 검색된 논문 중 해당 논문이 예를 들어 철학분야, 역사분야 등과 같이 본 논문의 분야에 무관하다고 판단된 부분에 대해 실시하였다.

검색식 번호	검색식	검색결과	
		논문	특허
		1987~2006	1977~2006
#1	TS=(constitution* or typolog*)	13,306편	6,807건
#2	#1 and TS=(Ayurveda or Ayurvedic or sasang or protein\$ or DNA or gene or genes or genetic or heritab* or epigenetic or SNP or informatic\$ or bioinformatic\$ or microarray or polymorphism\$ or polymorphic or genomic\$ or genomeor genotype\$ or phenotype\$ or haplotype\$ or proteomic\$ or physiome\$ or interactome\$ or pharmacogen* or diagnosis or diagnostic or prescription* or (personalized or customized)(w)medicine\$ or herbal or "Chinese medicine\$" or "oriental medicine\$" or "traditional medicine\$" or "natural product\$" or phytomedicine\$ or "ethnic group\$" or "chronic disease\$" or aging or ageing or epidemiolog* or "system biolog*")	4,633편	406건

【표 2】 검색 키워드 및 검색식

III. 본 론

검색 키워드에서 *와 \$ 표시는 와일드문자로서, word*는 word로 시작되는 모든 단어의 검색을 의미하며, word\$는 앞의 접두사에 상관없이 word로 끝나는 단어에 대한 검색조건을 의미한다.

논문검색에는 미국 Thomson ISI사에서 제공하는 학술정보 데이터베이스를 이용하였다.¹⁾ 이는 주요 과학기술저널 인용색인정보가 수록된 데이터베이스(DB)로, 총 6,431종의 우수한 과학기술 주요저널만을 DB로 구축하고 있어, 국제적인 인지도가 높을 뿐만 아니라 주요 국가, 연구기관, 저자들의 동향을 살피기에 적절하다.

한편, 특허검색에 이용된 Derwent사의 DWPI는 EPO, WIPO를 포함한 42개 주요국의 특허정보 약 1,400여만 건을 제공하는 대표적인 특허데이터베이스로서, 산업 전문분야에 걸쳐 방대한 특허를 제공하고 있으며, 특허패밀리 정보, Derwent 분류, 출원인 표준화 등 전문적인 특허정보를 제공하고 있다.

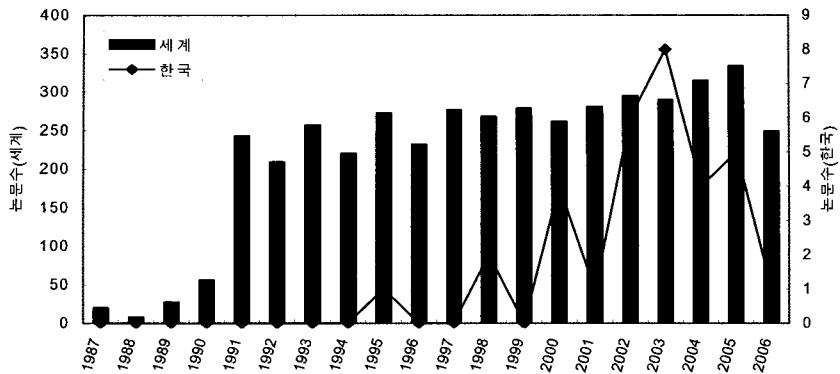
1. 논문 분석

1) 국내의 연구동향 분석

(1) 논문의 시계열 추이

지난 20년간 체질임상정보 및 바이오정보를 활용한 연구 논문 수는 연평균 13.4%의 비교적 높은 증가율로 꾸준한 증가추이를 보이고 있다(그림 1). 특히 인간게놈프로젝트(Human genome project; HGP)가 시작된 1990년을 기점으로 하여 1991년에는 4.3배의 폭발적인 논문수의 증가를 보이는 등 유전체정보 관련 연구 동향과 맥락을 거의 같이 하고 있으며, HGP 이전과 이후의 양상이 뚜렷이 구분되고 있다. 1990년대 및 2000년대를 거치면서는 매년 270여 편의 논문이 발표되면서 안정적인 증가추이를 보이고 있다.

1) Web of Knowledge(<http://portal.isiknowledge.com/>)
SCI-E(Science Citation Index -Expanded) (2007. 9. 7)



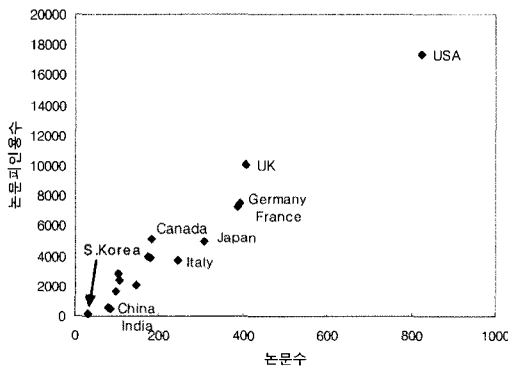
【그림 1】 체질임상정보 및 바이오정보 활용연구 논문 발표수 추이

번호	국가	논문수(편)	점유율(전체,%)	점유율(Top25,%)
1	USA	824	18.1	19.8
2	UK	407	8.9	9.8
3	France	391	8.6	9.4
4	Germany	386	8.5	9.3
5	Japan	306	6.7	7.4
6	Italy	245	5.4	5.9
7	Canada	182	4.0	4.4
8	Netherlands	180	3.9	4.3
9	Sweden	173	3.8	4.2
10	Spain	145	3.2	3.5
11	Australia	105	2.3	2.5
12	Switzerland	101	2.2	2.4
13	Belgium	97	2.1	2.3
14	China	85	1.9	2.0
15	India	80	1.8	1.9
16	Brazil	63	1.4	1.5
17	Austria	57	1.2	1.4
18	Denmark	51	1.1	1.2
19	Finland	50	1.1	1.2
20	Russia	50	1.1	1.2
21	Poland	41	0.9	1.0
22	Norway	38	0.8	0.9
23	Israel	36	0.8	0.9
24	South Korea	32	0.7	0.8
25	Turkey	28	0.6	0.7
-	기타	411	9.0	0.2
-	소계(Top25)	4,153	91.0	100.0
-	합계	4,564	100.0	-

【표 3】 체질임상정보 및 바이오정보 활용분야에 대한 각국의 논문발표 현황

(2) 국가 수준분석

저자 국적별로 학술논문 발표 현황을 통해 체질임상정보 및 바이오정보 활용 연구 분야에 대한 국가의 연구수준을 비교해 보면, 표 3과 같이 미국이 18.1%(824편)로 가장 큰 점유율을 보임으로써 기술 우위를 확고히 하고 있음을 알 수 있다. 미국의 뒤를 이어 영국, 프랑스, 독일, 일본의 순으로 각국이 6~8%의 점유율을 보임으로써 유럽강국 및 일본 또한 체질임상정보 및 바이오정보 활용연구를 활발히 진행하고 있음을 알 수 있다. 상위 25개국에 대한 점유율을 살펴본 결과도 전체 점유율의 결과와 유사한 결과를 보임으로써 상위국 몇몇 국가들에 의한 기술 점유의 영향이 큰 것으로 나타났다. 그러나 보통 과학기술분야에서 미국의 기술점유율이 30%를 넘는 것이 일반적인 현상이고, 유럽국들의 점유율도 10%를 넘는 경우가 많다는 점을 고려한다면, 본 연구 분야는 특정 국가에 의한 기술 지배력은 그리 높지 않은 것으로 보인다. 우리나라는 세계 24위이며 점유율에 있어서는 0.7%(32편)에 불과한 것으로 나타나 세계적으로 크게 뒤쳐져 있음을 알 수 있다. 결국 미국 및 일본을 위시하여 유럽 선진기술국들이 체질임상정보 및 바이오정보 활용연구를 주도하고 있다고 볼 수 있다.



【그림 2】 체질임상정보 및 바이오정보

활용분야에 대한 논문수와 논문피인용수의 상관관계

그림 2는 주요국들에 대해 논문수와 논문피인용수의 상관관계를 나타낸 것으로, 논문수와 논문 피인

용 횟수와의 사이에 양의 상관관계가 성립하고 있음을 알 수 있다. 논문의 피인용수가 논문의 질(質)을 일부 반영한다는 전제 하에서 볼 때 미국을 위시하여 영국, 독일, 프랑스, 일본, 캐나다 등이 중국이나 인도, 우리나라에 비해 논문 수뿐만 아니라 논문 피인용수도 높게 나타나는 이러한 결과는 논문의 영향력 또한 세계 수준에 크게 뒤쳐져 있는 현실을 반영한다는 점에서 시사하는 바가 크다.

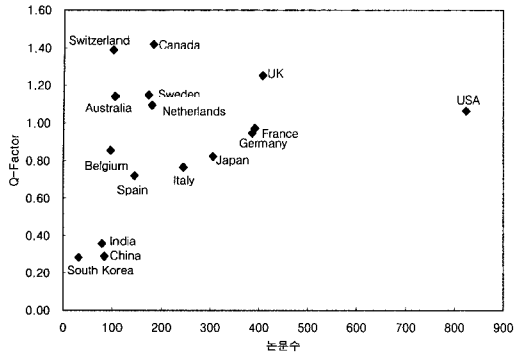
한편, 우리나라의 연구수준을 보다 객관적인 지표로 판단하기 위해 연구성과평가 전문기관인 네덜란드 Leiden대학의 CWTS(Center for Science & Technology Studies)에서 개발한 분석지표²⁾를 일부 수정한 Q-Factor를 이용하여 분석하였으며, 그 결과를 그림 3에 나타내었다(각 지표에 대한 상세설명은 표 4 참조).

분석 지표	지표 내용
P(A,F)	국가 A가 분야 F에 발표한 논문의 총 수
C(A,F)	국가 A가 분야 F에 발표한 논문의 총 피인용 수
CPP(A,F)	국가 A가 분야 F에 발표한 논문 1편당 피인용 수 = C(A,F)/P(A,F)
CPP(W,F)	분야 F에 발표된 논문의 세계 평균 피인용 수 (분야별 평균 피인용 수)
Q-Factor Q(A,F)	분야 F에 발표된 논문의 평균 피인용 수에 대한 국가 A의 해당 분야 논문의 평균 피인용 수(수준 평가 지표) = CPP(A,F)/CPP(W,F)

【표 4】 피인용수 기반의 논문 성과 분석용 지표³⁾

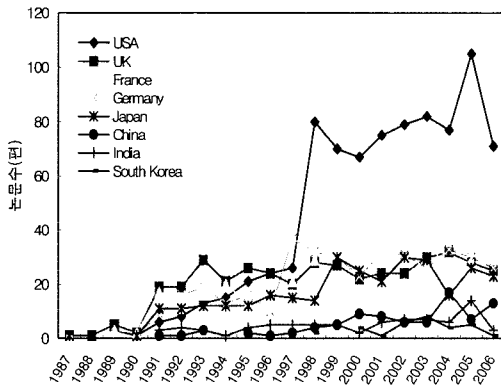
Q-Factor는 평가 대상 국가가 발표한 논문 1편당 피인용수[CPP(A,F)]와 해당 분야의 세계 평균 피인용수[CPP(W,F)]의 비율을 나타내는 것으로, Q-Factor가 1인 경우 평가대상에 의한 논문의 평균 피인용 수가 세계 수준과 같음을 의미하며, 1보다 큰 경우 세계 평균 수준보다 우수함을 의미한다.

2) A.F.J. van Raan and Th. N van Leeuwen. Assessment of the Scientific Basis of Interdisciplinary, Applied Research. CWTS, University of Leiden. 2002. 31. pp. 611-632
 3) 이상필. 연구성과의 질 제고를 위한 논문평가 모형개발. 기술혁신학회지. 2005. 9(3). pp. 538-557.

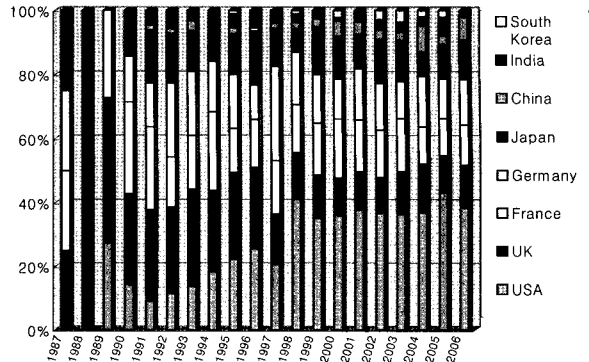


【그림 3】 체질임상정보 및 바이오정보 활용분야 논문 성과물의 국가 활동도

각 국가들의 연구수준이 시대의 흐름에 따라 어떻게 변화하였는지 알아보기 위해 상위5개국 및 우리나라, 중국, 인도의 지난 20년간 논문 수 발표추이를 살펴보면, 그림 4 및 그림 5와 같이 1990년대 후반 미국의 약진이 돋보인다. 최근 10년간 논문수의 연평균증가율에 있어서는 미국이 10.6%, 일본이 4.4%로 프랑스, 영국, 독일의 유럽국에 비해 꾸준한 성장을 하여 왔으며, 특히 연평균 20.6%의 가장 높은 증가율을 보인 중국의 최근의 급증세가 두드러진다. 우리나라도 꾸준한 증가추이를 보이고는 있으나 전체적인 활동도가 미미하여 통계적인 의미를 두기는 힘들다. 하지만 전체적으로 아시아 국가들이 최근에 연구 확대가 두드러지게 나타나고 있다.



【그림 4】 주요국의 체질임상정보 및 바이오정보 활용분야 논문 발표추이



【그림 5】 주요국의 체질임상정보 및 바이오정보 활용분야 논문 점유율 변화

5년의 간격으로 체질임상정보 및 바이오정보 활용분야 논문 수에 따른 국가순위를 나타낸 표 5에서 미국, 독일, 프랑스, 영국, 일본 등 기술 강국들의 상위랭킹에는 큰 변화가 없으나 우리나라는 43위→28위→21위로 약진한 결과를 보이고 있다. 비록 앞에서도 살펴본 바와 같이 체질임상정보 및 바이오정보 활용분야에 있어서 그동안 우리나라의 연구 성과가 아직까지는 양적·질적으로 미미하기는 하지만 향후 많은 발전가능성을 내포하고 있음을 알 수 있다.

체질임상정보 및 바이오정보를 활용한 연구동향분석에 대한 연구

번호	T1(1987~1991)		T2(1992~1996)		T3(1997~2001)		T4(2002~2006)	
	국가	논문수	국가	논문수	국가	논문수	국가	논문수
1	UK	28	UK	119	USA	318	USA	414
2	France	23	USA	81	Germany	154	Germany	146
3	Japan	14	France	74	France	148	France	144
4	Canada	14	Germany	74	UK	121	UK	139
5	Germany	11	Japan	63	Japan	105	Japan	124
6	USA	10	Italy	55	Italy	86	Italy	95
7	Netherlands	10	Sweden	47	Netherlands	68	Canada	76
8	Italy	8	Canada	43	Sweden	66	Netherlands	60
9	Sweden	7	Netherlands	42	Spain	64	Spain	58
10	Australia	6	Switzerland	27	Canada	49	Sweden	53
11	Spain	5	Australia	24	Australia	36	China	49
12	Switzerland	5	Belgium	20	Switzerland	31	Belgium	44
13	Russia	5	Spain	18	Belgium	31	Australia	39
14	India	3	India	17	China	28	Switzerland	38
15	Brazil	3	Denmark	13	Brazil	27	India	37
16	Norway	3	Norway	11	India	23	Brazil	28
17	Taiwan	3	Israel	11	Austria	20	Finland	26
18	Belgium	2	Austria	10	Denmark	16	Austria	25
19	Austria	2	Finland	10	Russia	16	Russia	25
20	Argentina	2	China	7	Poland	14	Poland	25
21	China	1	New Zealand	7	Israel	14	South Korea	24
22	Denmark	1	Taiwan	6	Finland	13	Denmark	21
23	Finland	1	Greece	6	Norway	13	Turkey	17
24	Israel	1	Brazil	5	Czech Rep.	9	Argentina	16
25	Turkey	1	Hong Kong	5	Turkey	8	Mexico	14
26	Hong Kong	1	Russia	4	New Zealand	8	Taiwan	12
27	Venezuela	1	Mexico	4	Ireland	8	Greece	12
28	Iceland	1	Argentina	3	South Korea	7	Norway	11
29	Algeria	1	Czech Rep.	3	Croatia	7	Czech Rep.	11
30	NAMIBIA	1	Ukraine	3	Egypt	7	Portugal	11
31	Poland	0	Poland	2	Argentina	6	Israel	10
32	South Korea	0	Turkey	2	Taiwan	6	New Zealand	9
33	New Zealand	0	Portugal	2	Portugal	6	Croatia	9
34	Czech Rep.	0	Croatia	2	Singapore	6	Chile	7
35	Mexico	0	Hungary	2	Mexico	5	Hungary	6
36	Greece	0	Egypt	2	Greece	5	Iran	6
37	Portugal	0	Chile	2	Hungary	5	Ireland	5
38	Croatia	0	Saudi Arabia	2	Bulgaria	5	Venezuela	5
39	Hungary	0	Kuwait	2	Saudi Arabia	4	Tunisia	5
40	Ireland	0	Yugoslavia	2	Yugoslavia	4	Slovenia	4
41	Egypt	0	Slovakia	2	Hong Kong	3	Egypt	3
42	Chile	0	Czechoslovakia	2	Venezuela	3	Singapore	3
43	Singapore	0	South Korea	1	Kuwait	3	Lebanon	3

【표 5】 체질임상정보 및 바이오정보 활용분야 논문 수에 따른 국가랭킹 추이

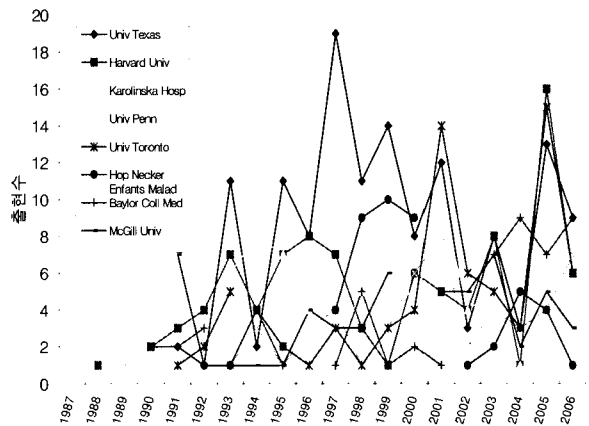
(3) 연구기관 분석

지난 20년간 체질임상정보 및 바이오정보 활용분야 논문을 가장 많이 발표한 연구기관은 표 6과 같이 1.4%의 점유율을 나타낸 미국 텍사스대학의 뒤를 이어, 미국 하버드대학, 캐나다 토론토대학, 스웨덴의 카롤린스카 병원, 미국 펜실베니아대학의 순으로 나타났다. 상위 20위 기관에 대한 점유율로 봤을 때에도 가장 출현수가 많았던 미국 텍사스대학의 점유율이 10%에 못 미치는 것으로 나타나 체질 및 바이오정보 활용분야에 대해 집중적인 연구 활동을 펼치는 연구기관은 보이지 않으며, 수많은 대학과 병원, 연구소가 관여하고 있음을 알 수 있다. 다만, 유전체정보나 암유전자 관련 연구에 주력하는 세계적인 대학 및 병원의 활동이 약간 두드러지는 것으로 보인다.

번호	연구기관	출현수	점유율(전체,%)	점유율(Top30,%)
1	Univ Texas	137	1.4	9.8
2	Harvard Univ	93	0.9	6.7
3	Univ Toronto	71	0.7	5.1
4	Karolinska Hosp	65	0.7	4.7
5	Univ Penn	55	0.6	3.9
6	McGill Univ	54	0.5	3.9
7	Hop Necker Enfants Malad	53	0.5	3.8
8	Baylor Coll Med	52	0.5	3.7
9	Karolinska Inst	47	0.5	3.4
10	NCI	47	0.5	3.4
11	Univ Pittsburgh	44	0.4	3.2
12	Univ Washington	44	0.4	3.2
13	Childrens Hosp Philadelphia	43	0.4	3.1
14	Childrens Hosp	41	0.4	2.9
15	Hosp Sick Children	38	0.4	2.7
16	Univ So Calif	38	0.4	2.7
17	Univ Helsinki	37	0.4	2.7
18	Univ Michigan	37	0.4	2.7
19	Washington Univ	37	0.4	2.7
20	Univ Amsterdam	36	0.4	2.6
-	기타	8882	89.3	
소계(Top20)		1069	10.7	100.0
합계		9951	100.0	

【표 6】 체질임상정보 및 바이오정보 활용분야의 세계 주요 연구기관 현황

또한 이들 주요 연구기관의 활동을 시계열 추이를 통해 살펴보면, 그림 6과 같이 대부분의 주요기관이 매년 10편 내외의 비교적 소수의 논문을 발표하고 있어 통계적인 의의를 두기는 힘들다. 다만, 지난 20년을 통해 타 기관에 비해 월등한 연구 성과를 보이는 기관은 찾아보기 힘들며, 이는 본 연구 분야가 아직 태동기에 있으며, 탐색단계에 머물러있는 데 기인하는 것으로 파악된다.



【그림 6】 체질임상정보 및 바이오정보 활용분야 세계 주요 연구기관의 논문발표추이

한편, 우리나라의 경우는 표 7에서 보는 바와 같이 원광대, 경희대 등 한의과대학을 중심으로 체질임상정보 및 바이오정보 활용 관련 연구를 일부 수행하는 것으로 나타났으며, 전반적으로는 연구가 그다지 활성화 되어 있지 못한 상황이다.

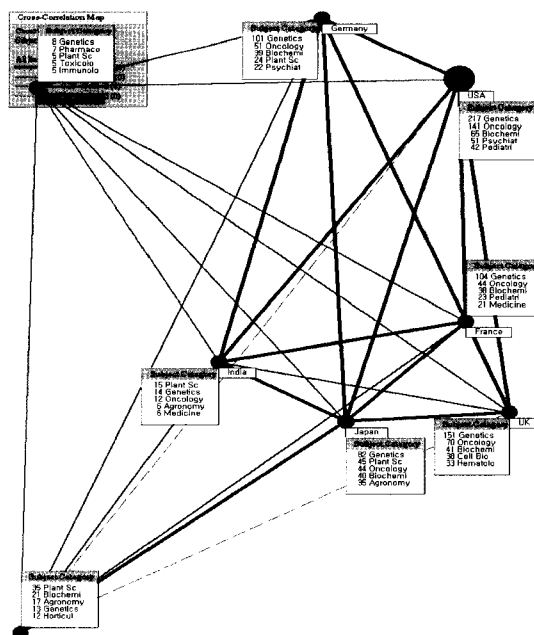
연구기관	논문수(편) ⁴⁾	발생수(편)	논문수 비중(%)	발생수 비중(%)
Wonkwang Univ	11	28	18.0	31.8
Kyunghee Univ	10	12	16.4	13.6
Chonbuk Natl Univ	3	4	4.9	4.5
Seoul Natl Univ	3	5	4.9	5.7
Cheju Natl Univ	2	2	3.3	2.3
Chungnam Natl Univ	2	2	3.3	2.3
Dongguk Univ	2	2	3.3	2.3
Jeonbuk Natl Univ	2	2	3.3	2.3
Kangweon Natl Univ	2	2	3.3	2.3
Kyungpook Natl Univ	2	3	3.3	3.4
Woosuk Univ	2	2	3.3	2.3
기타	20	24	32.8	27.3
합계	61	88	100	100

【표 7】 체질임상정보 및 바이오정보 활용 관련 국내 주요연구기관

2) 연구분야 분석

(1) 국가/연구분야 상관분석

그림 7은 데이터 셋의 주제분야(subject category)를 교차상관(cross-correlation) 변수로 하여 국가 간 네트워크 현황을 살펴본 것이다. 체질임상정보 및 바이오정보 활용 연구 활동이 비교적 활발한 상위 국가를 중심으로 살펴본 결과, 가장 활발한 연구 활동을 펼치고 있는 미국과 유럽3국(프랑스, 독일, 영국), 일본, 인도는 유전학, 종양학, 생화학 분야 등 첨단 생명공학 기술 분야에서 강한 네트워크를 형성하고 있는 반면, 우리나라와 중국은 이들 국가들과 연구 분야 공유도가 상대적으로 약하며, 우리나라를 비롯한 중국, 인도, 일본 등 아시아국들은 한약이나 생약에 관한 식물과학 분야에서 비교적 강한 네트워크를 형성하는 것으로 나타나 국가마다 주력 연구 분야를 달리 하고 있음을 알 수 있다.

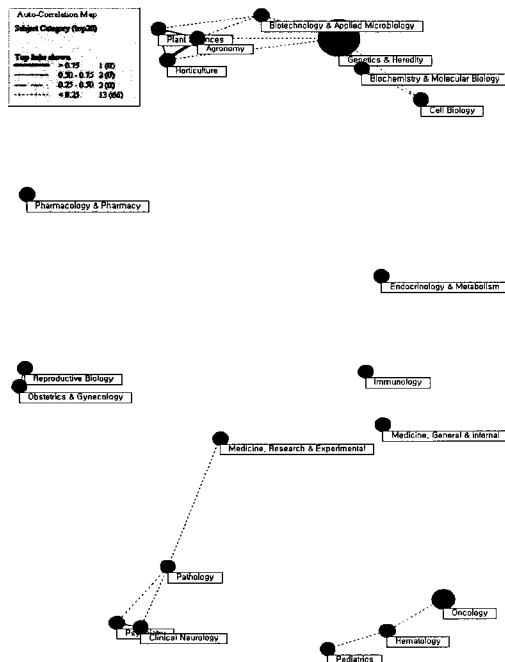


【그림 7】 주제분야(subject category)를 교차상관변수로 한 경우의 국가네트워크 현황

4) 동일 논문을 2개 이상의 연구기관이 공동연구로 저술하였을 경우 연구기관별로 한편씩의 논문을 부여하여 논문 수를 산출하므로 연구기관별 논문 수의 합은 분석모집단의 총 논문 수와 일치하지 않음.

(2) 연구 분야간 상관분석

체질임상정보 및 바이오정보 활용에 관한 연구 분야를 분석하기 위해 데이터 셋의 주제분야(subject category)로 자기상관분석(auto-correlation analysis)을 실시한 결과, 그림 8과 같이 식물과학, 농학 및 육종학이 강한 네트워크를 형성하고 있고, 여기에 유전학, 생명공학 및 응용미생물, 생화학 및 분자생물, 세포공학 등 첨단 생물공학기술이 링크를 형성하면서 가장 큰 클러스터를 형성하고 있음을 알 수 있다. 또한 병리학 등 기초의학과 정신의학, 임상신경학이, 그리고 소아과학, 혈액학, 종양학이 보다 강한 상관성을 보이며 연구영역을 형성하고 있는 것으로 보인다. 체질임상정보 및 바이오정보 활용 연구는 결국에는 이러한 기초적인 임상연구 자료를 통해 이를 진단하고 이를 치료하기 위해 천연물(한약이나 생약)을 활용하고자 하는 목적을 가지고 있으므로, 이러한 제반연구에 네트워크상에 나타난 다양한 학제의 연구 분야가 동원되어야 할 것으로 판단된다.



【그림 8】 체질임상정보 및 바이오정보 활용연구 관련 연구 분야간 네트워크 현황

(3) 키워드로 본 연구 분야 분석

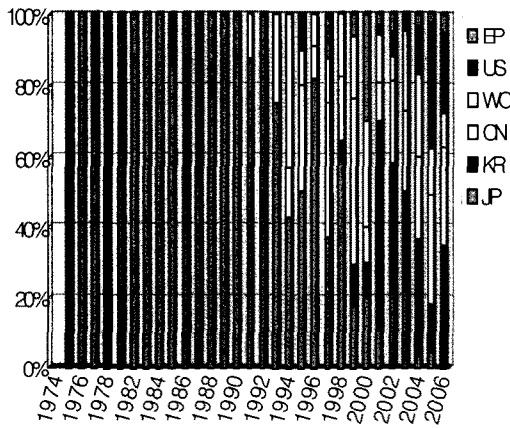
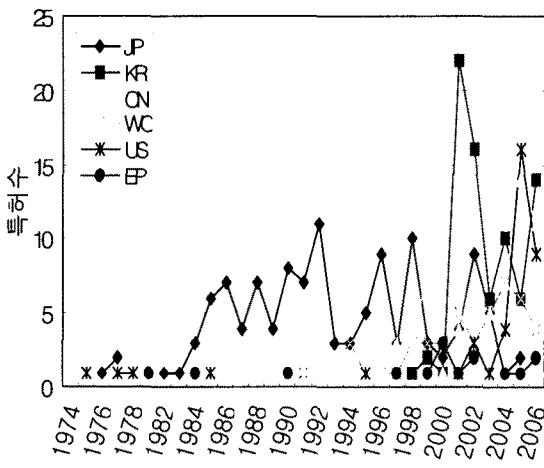
그림 9는 체질임상정보 및 바이오정보 활용연구와 관련 키워드를 분석한 결과이다. 구간별 다출 키워드 50개를 중심으로 동시발생분석(co-occurrence analysis)을 통해 단어(키워드) 간 네트워크 현황을 살펴보고 시대별 추이를 통해 연구 분야의 흐름을 유추해 본 것이다.

전체적으로는 암이나 암유전자, 유전자 변이로 이루어진 그룹과, 세포유전학 및 염색체 이상으로 이루어진 또 다른 그룹의 두 개의 기술 군을 가지는 것으로 나타났다. 체질은 후자의 그룹, 특히 염색체 이상과 강한 상관성을 보였다.

키워드 출현 추이변화를 통해 기술의 흐름을 살펴보면, 초창기에는 암이나 암유전자 등의 정보가 유전자 클론, 하이브리드, 유전자자리 등 유전자 조작과 관련하여 강한 연관성을 가짐으로써 전체적으로 광범위한 유전자 정보 클러스터를 형성하고 있음을 알 수 있다. 점차 암이나 암유전자 등의 키워드와 염색체와 관련된 유전자조작 관련 키워드 간의 링크가 강화되고, 유전체 및 세포유전학이 새로이 등장하기 시작하였으며, 암과 유전자변이 간의 관련성이 깊어지고, 염색체 이상을 비롯한 세포유전학적인 연구가 점차 큰 클러스터를 형성하기 시작하였음을 확인할 수 있다.

이후에 암이나 암유전자, 유전자변이 관련 연구는 링크 강도가 약해지고, 유전체, 염색체이상, 세포유전학이 하나의 큰 군집을 형성하였다. 질병 연구를 대상 군에 대한 탐색에서 시작하는 유병 율이나 위험인자, 생물다양성 및 발병집단 등에 대한 연구가 별도의 클러스터를 형성하는 것이 눈에 띈다. 결국 체질정보는 염색체 이상과 어느 정도의 연관성을 보이고는 있으나 전반적으로는 유전자 관련 키워드들과 관련도가 적게 나타나고, 여타 키워드들과 링크를 잘 형성하지 못하는 점으로 보아 체질정보를 유전정보와 같은 하나의 바이오정보로 활용하는 연구는 아직 구심점을 형성하지 못한 연구의 태동단계로 판단된다. 키워드 클러스터의 시대적 추이를 분석해 보면, 최근에 들어서야 역학, 유병율, 생물다양성 등 질병의 대상 집단에 대한 심층적인 연구가 유전체정보

그림 12는 출원국에 따른 특허출원 동향을 살펴본 것으로, 일본특허는 거의 변화가 없는 반면, 우리나라, 미국, 중국은 2000년대 들어 출원이 증가하는 경향을 보임으로써 논문의 국가별 증가 추이와 유사한 형태를 보인다. 특히 최근 들어 국내출원인의 활동이 두드러진 것으로 나타나 국내의 경우 기초연구 활동과는 달리 체질임상정보나 바이오정보의 활용 및 응용 기술개발 움직임이 일고 있으며, 이러한 활동이 특허출원으로 이어지고 있는 것으로 사료된다.



【그림 12】 체질임상정보 및 바이오정보 활용 관련 주요국 특허출원 추이 및 점유율 변화

(3) 연구기관 분석

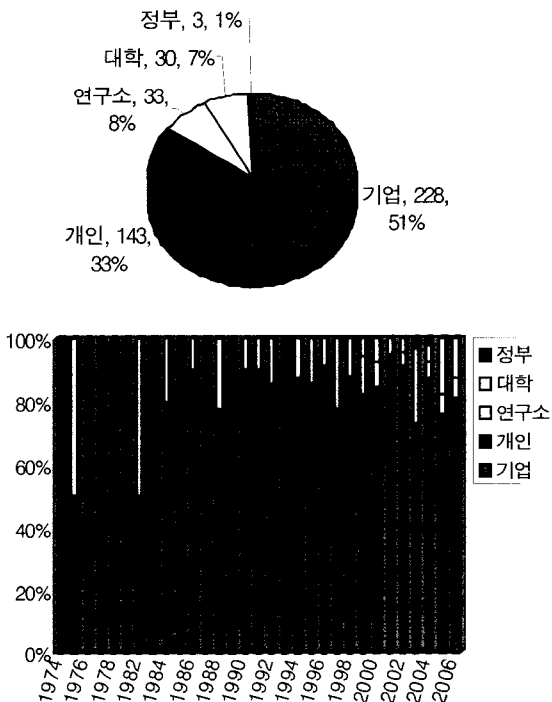
표 8은 지난 30여 년간 체질임상정보 및 바이오정보 활용 관련 특허출원 활동이 활발한 주요 연구기관과 발명자를 나타낸 것이다. 출원인(연구기관)이나 발명자(연구자) 모두 우리나라의 개인출원(Suh Y H)이 두드러지게 나타났으며, 최다 출원인이나 발명자의 비중이 전체에 비해 극히 미미하게 나타남으로써 본 기술영역이 특정 연구기관이나 연구자에 의해 주도되지는 않고 있음을 알 수 있다. 또한 주요 출원인의 대부분이 일본의 영상이나 측정기기 관련 대기업들로 체질임상정보나 바이오정보를 측정하는 기구나 장치와 관련된 기술을 보유한 기업들의 활동이 돋보인다. 한편, 발명자들은 한국, 중국 국적이 많아 이들이 한의학이나 한약에 대한 경험을 토대로 개인 혹은 기업 활동을 통해 기술개발에 참여하고 있는 것으로 사료된다.

번호	주요 출원인	특허수	비중 (%)	주요 발명자	특허수	비중 (%)
1	SUH Y H (한국)	16	2.9	SUH Y H (한국)	16	2.5
2	FUJI PHOTO FILM CO LTD (일본)	14	2.5	FILIMONOV S N (러시아)	4	0.6
3	KONICA CORP (일본)	8	1.4	JUNG H N (한국)	4	0.6
4	MITSUI CHEM INC(일본)	6	1.1	WANG Y (중국)	4	0.6
5	HITACHI LTD (일본)	5	0.9	KIM K (한국)	3	0.5
6	TORAY IND INC (일본)	5	0.9	KIM W G (한국)	3	0.5
7	TOSHIBA KK (일본)	5	0.9	OU X (중국)	3	0.5
8	UNIV FUDAN (중국)	5	0.9	TU Y (중국)	3	0.5
9	JUNG H N (한국)	4	0.7	XU H (미국)	3	0.5
10	CANON KK (일본)	3	0.5	ZHU N (중국)	3	0.5
11	INST PASTEUR(프랑스)	3	0.5			
12	KIM K (한국)	3	0.5			
13	MELJI MILK PROD CO LTD (일본)	3	0.5			
14	NOVOK DOCTORS TRAINING INST(러시아)	3	0.5			
15	OLYMPUS OPTICAL CO LTD (일본)	3	0.5			
16	SUMITOMO CHEM IND KK (일본)	3	0.5			

【표 8】 체질임상정보 및 바이오정보 활용 관련 주요 출원인 및 발명자

5)WO: World patent;EP: European patent

출원인의 유형을 기업, 연구소, 대학, 정부, 개인으로 나눠 연구기관의 성격을 구분하여 분석한 결과, 그림 13과 같이 기업이 절반을 넘고 개인이 33%의 높은 점유율을 보여, 대학이나 연구소의 활동이 높았던 논문 분석 결과와는 차이가 있으며, 이는 논문발표나 특허출원에 대한 연구기관의 일반적인 전략적 성향에 따르는 것으로 풀이된다. 최근 들어 기업의 출원보다는 개인 출원이 증가하는 현상은 과거 체질 및 바이오 정보의 측정을 위한 기기회사의 기술개발 활동 중심에서 체질 및 바이오 정보를 진단이나 치료를 위해 의학 분야에 접목하여 활용하는 기술개발로 옮겨가고 있으며, 이러한 활동이 개인적 차원의 특허출원으로 다양하게 나타나고 있기 때문인 것으로 추측된다.



【그림 13】 체질임상정보 및 바이오정보 활용 관련 출원인 유형별 점유율 추이

2) 연구 분야 분석

(1) 국가/연구 분야 상관분석

그림 14는 출원국가별 주요 기술의 프로파일을 나타낸 것이다. 주요 기술군은 Derwent classification(DC)에 따라 전체 데이터에서 점유율 1.5% 이상을 차지하는 다출원 상위 기술군 12개로 한정하였다. DC는 특허에서 다루는 기술의 내용을 표현하기 위한 분류 체계이다. 일반적으로 사용되는 국제특허분류(IPC⁶⁾, International Patent Classification)가 심사관 또는 국가에 따라 다르게 부여될 수 있다는 점에 반하여 DWPI⁷⁾ 데이터베이스의 DC는 자체 전문가에 의하여 일관되게 부여된다는 특징이 있다.

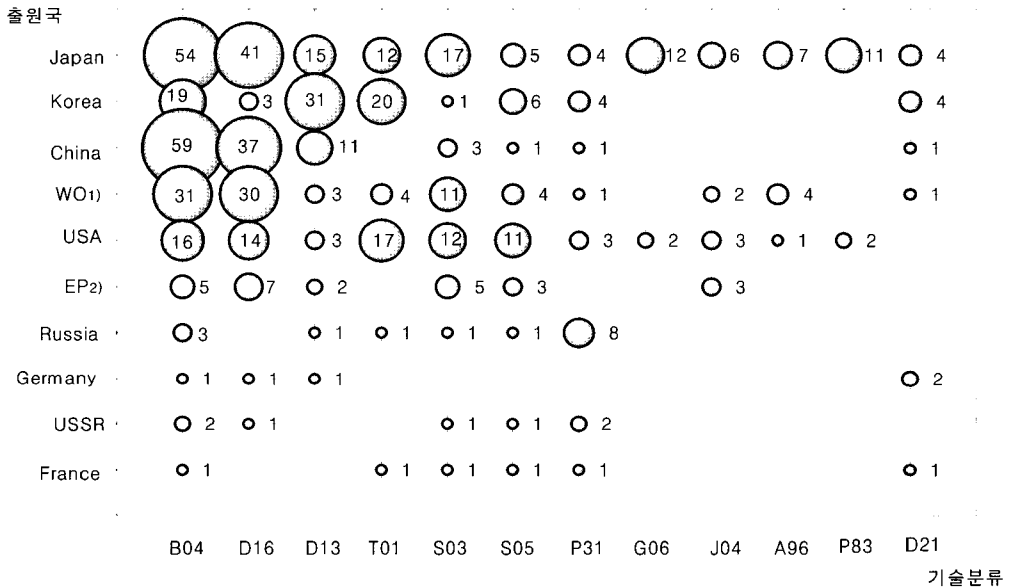
DC 분류체계의 가장 상위에는 화학, 일반 공학 및 전기·전자공학의 3개 기술 군(Technology Group)이 존재하며, 기술 군 하부에는 전체 21개의 섹션이 존재한다. 각 섹션은 A-M(화학), P-Q(일반 공학), S-X(전기·전자공학)로 구성된다. 섹션의 하부에는 291개의 세부 분류인 클래스(Class)가 존재하며, 화학 군에는 138개, 일반 공학에는 103개, 전기·전자공학에는 50개의 클래스가 있다. 예를 들어 X22는 Automotive Electrics에 해당하는 클래스이며, 이는 섹션 X인 Electric Power Engineering의 하부에 존재한다. 본 분석에서는 분석 범위를 클래스 단위로 한정하였다.

체질임상정보 및 바이오정보 활용에 관한 특허출원이 많은 것으로 나타난 일본, 우리나라, 중국, 미국의 경우에 있어서 우리나라를 제외하고는 B04(천연물 및 고분자) 및 D16(유전공학 등)의 비중이 높게 나타나 한의학적 측면에서 체질 및 바이오 정보를 활용하는 기술개발의 움직임이 있음을 확인할 수 있는 반면, 우리나라의 경우는 B04 및 D16 보다도 D13(식품, 사료 처리) 분류가 가장 많이 나타나는 특징을 보였다. 이는 천연물에서 유래하는 건강식품류에 관해 특히 개인출원의 형태로 특허출원이 많이 행해지

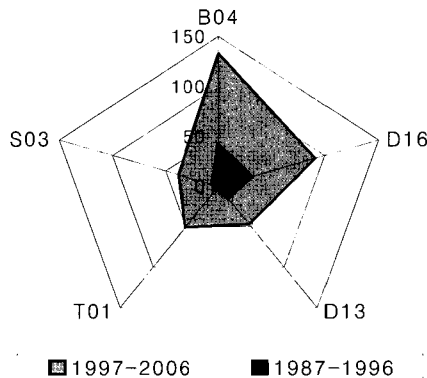
6) 현재 특허문헌에 대해 전세계적으로 통용되고 있는 분류 체계로 제8판이 이용되고 있다. 기술전체를 8개의 section으로 나누고, 각각에 대해 class, subclass, group, subgroup으로 세부화 하였다.
7) DWPI에서는 클래스 하부에 5단계의 세부 기술 분류인 메뉴얼 코드를 두고 있다.

고 있는 우리나라의 특수한 경우를 반영하고 있기 때문인 것으로 풀이된다. 또한 T01(컴퓨터 및 제어) 분류가 타 국가에 비해 많은 부분을 차지하는 것으로 나타난 점은 IT 기술 분야에 강점을 지닌 우리나라의 특성을 반영하는 것으로 보인다. 체질임상정보 및 바이오정보의 활용연구 분야에 있어서 국가 혹은 연구 집단의 특성과 강점을 살린 연구 분야 개척과 개발이 가능할 것으로 사료된다.

그림 15에는 DC 분류체계에 따른 5대 주요 기술 분야의 과거 10년과 최근 10년의 특허출원 동향을 나타내었다. 우리나라가 특히 강한 면을 보인 T01(컴퓨터 및 제어) 분야가 4건에서 51건으로 12.8배의 가장 높은 성장을 보였으며, S03(광학측정기기 등)은 6.3배, B04(천연물 및 고분자), D16(유전공학), D13(식품, 사료 처리)은 약 3배의 성장에 그쳐 바이오정보 측정을 위한 컴퓨터 및 제어기술의 발전 속도가 가속화되고 있음을 알 수 있다.



【그림 14】 체질임상정보 및 바이오정보 활용 관련 특허출원 국별 기술 분류(DC) 프로파일



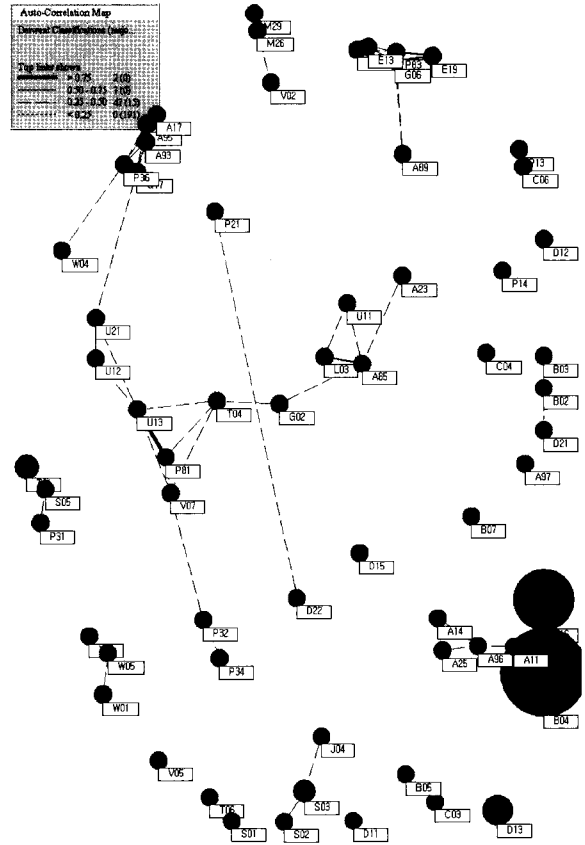
【그림 15】 체질임상정보 및 바이오정보 활용 관련 기술 분야 추이

(2) 연구 분야간 상관분석

체질임상정보 및 바이오정보 활용에 관한 연구 분야간 관계를 분석하기 위해 DC 분류체계를 이용하여 자기상관분석(auto-correlation analysis)을 실시한 결과, 그림 16과 같이 전체적으로는 B04(천연물 및 고분자)와 D16(유전공학)이 주류를 나타내었다. 이는 대부분 천연 고분자로 이루어진 한약이나 생약 정보 등 한의학 분야에서 체질 및 바이오 정보를 활용하고자 하는 본 기술의 적용범위에 기인한 것으로 보인다. 또한 반도체 및 전자회로(U12, U13 및 U21, U11 등), 인쇄 및 광학기술(G02, G06 등), 건강관련 제품(P31, P32, P34, P36, P83 등), 컴퓨터 및 제어(T01 등), 측정기구(S05, S03, S02 등), 통신분야(W01, W05 등), 유기화합물(E13, E19 등) 및 고분자(A14, A25, A11, A95, A96, A93, A89 등) 등 유전정보를 측정하기 위한 장치나 도구, 단백질 등 유기화합물이나 고분자 관련 기술 등 많은 종류의 기술들이 상호 연관성을 가지며 다양한 크기의 네트워크를 형성하고 있어 체질 및 바이오 정보 활용연구가 생명공학, 기계 및 전자통신, 정보공학, 유기화학 등 다양한 분야의 기술영역을 포함하는 융복합기술임을 알 수 있다.

체질 및 바이오 정보 활용연구와 관련된 기술들이 시대적으로 어떻게 변화해 왔는가를 알아보기 위해 Derwent 분류체계(DC)를 이용하여 기술 분류간 네트워크 형성과정을 시대적 추이에 따라 살펴본 결과, 전 구간을 통해 B04(천연물 및 고분자)와 D16(유전공학)이 핵심을 이루면서 점차 관련 분류체계가 많아지고 복잡해졌음을 알 수 있다. 1991년 이전까지는 B04(천연물 및 고분자)와 D16(유전공학)을 주축으로 광학측정기구(S03)나 진단장치(J04), 고분자 및 단백질(C03) 관련 기술이 영역을 이루고 있으며, 1980년대 후반 들어 본 연구영역에 등장하기 시작한 컴퓨터 및 제어(T01)와 통신 관련 기술(W01)이 1990년대 중반을 거치면서 보다 확고한 영역을 형성하기 시작한 것으로 보인다. 별도로 의료진단장비와 관련된 영상기기분야(G06, P83, A89 등)도 일정한 영역을 형성하고 있음을 알 수 있다. 최근(2002~2006년)에는 B04(천연물 및 고분자)가 가장 큰 영

역을 구축하고 있는 가운데 관련 기술영역들이 다양하게 출현하고 있고 네트워크의 밀도도 조밀해졌다.



【그림 16】 체질임상정보 및 바이오정보 활용 관련 특허 기술 분야간 네트워크 현황

IV. 결론

논문과 특허라는 정보분석을 통해 체질임상정보 및 바이오정보를 활용한 연구에 대한 분석한 결과는 국내외 모두 이제 연구가 태동되어 시작하는 단계로 정의할 수 있어 연구의 진입단계에서 충분히 경쟁력을 가지고 있으며, 최근 들어서야 활발히 연구가 진행되는 추이를 반영하면 성장가능성이 높은 연구라고 판단되어 진다.

논문 분석의 결과 중에서 유의한 점은 다음과 같다.

첫째, 논문저자의 국적이 한국인 논문의 시계열 추이분석을 통해 우리나라의 연구동향을 살펴보면, 2000년대 들어서야 매년 5편 내외의 논문이 발표되는 등 저조한 움직임을 보이고 있음을 알 수 있다. 지금까지 총 32편에 그친 저조한 연구 활동으로 인해 우리나라의 연구동향을 파악하기는 힘들며, 실제적으로 체질이나 바이오 정보를 활용한 연구는 지금까지 전무하다고 할 수 있다.

둘째 분석 대상 연구 분야에 대한 국가들에서 발간된 논문들의 점유율을 분석한 결과 본 연구 분야는 특정 국가에 의한 기술 지배력은 그리 높지 않은 것으로 보인다.

셋째 연구논문 발표 추이를 분석한 결과 전체적으로는 미국이나 유럽강국들이 연구 성과를 많이 내고 있음은 분명하나 우리나라를 위시하여 일본, 중국, 인도 등 아시아국들의 최근의 약진이 점차 두드러지고 있어 향후 많은 성장잠재력을 지니고 있음을 알 수 있다.

특허분석 결과 우리나라도 2000년대 들어서 특허 출원수가 급증하면서 34.1%의 높은 연평균증가율을 보이고 있으며, 이러한 점은 학술연구 활동이 미비한 논문분석 결과와는 다소 대비되는 부분이다. 우리나라의 경우 체질임상정보 및 바이오정보에 대한 학술적 연구보다는 한의학적 활용연구에 보다 주력하는 것으로 보인다.

또한 특허 관련되어 연구 분야 상관분석 결과 컴퓨터 및 제어기술이 전자진단장치와 강한 네트워크를 구축하면서 성장한 것으로 나타난 점은 체질 및 바이오 정보 활용분야에서 생물정보학(Bioinformatics) 등 새로운 영역의 등장과 성장을 반영하고 있는 것으로 보이며, 체질, 유전 등의 바이오 정보와 이를 효율적으로 처리하기 위한 컴퓨팅 기술의 접목은 향후 본 연구의 흐름을 주도할 수 있는 핵심기술로 여겨진다.

모형개발. 기술혁신학회지. 2005. 9(3).

2. A.F.J. van Raan and Th. N van Leeuwen. Assessment of the Scientific Basis of Interdisciplinary, Applied Research. CWTS, University of Leiden. 2002.

<전자문서>

1. Web of Knowledge (<http://portal.isiknowledge.com/>) SCI-E(ScienceCitation Index -Expanded)

참고문헌

<논문>

1. 이상필. 연구성과의 질 제고를 위한 논문평가