

The Analysis and Visualization of Interdisciplinary Characteristics

A Case Study of the International Epigenetics

Hongfang Shao

School of Public Health
Shanxi Medical University
Taiyuan, China
shftanya@126.com

Zhiguang Duan*

School of Public Health
Shanxi Medical University
Taiyuan, China
dzg52827@aliyun.com

Abstract—In order to demonstrate the interdisciplinary features, the authors retrieved ‘epigenetic*’ in their title, abstract or keywords from Web of Science between 1970 and 2012 to be the dataset of a case study, and used scientific metrology methods and information visualization tools to perform the discipline co-occurrence analysis and draw the interdisciplinary co-occurrence maps, which shows the interdisciplinary features and evolving trends in each periods during last four decades, the relationship and the association degree between disciplines, the core discipline clusters and the relative clusters. Visual analytics of the disciplines provides a valuable, timely and flexible approach so as to track the discipline development process and analyze the particularity in interdisciplinary development, which we expect will play a more active role in supplement to traditional research methods of discipline.

Keywords—interdisciplinary; visualization; mapping knowledge domains; co-citation analysis; epigenetics

I. INTRODUCTION

Interdisciplinary is ‘inter-discipline’ or ‘interdisciplinary’ research activities, which results in a knowledge system of cross-disciplinary science[1]. A discipline research and its numerous disciplines involved lies in a relationship of cross-research, sucks nutrition from various disciplines related to, and constantly improves itself and develops in the cross-development, which is a reflection of discipline to be mature[2]. Therefore, interdisciplinary characteristics research is an effective viewing angle to detect whether a discipline is mature. Using scientific metrology methods and information visualization techniques, we could make the interdisciplinary characteristics better embodied and achieved [3].

Epigenetics, a subdiscipline of genetics studying epigenetic variation, is a rapidly growing and fast-moving interdisciplinary field of study, revealing a large cross-discipline and a large fusion development trend, involving the genetic gene expression changes without DNA sequence changes [4,5]. As the research features of the diversity and openness in epigenetic interdisciplinary study, we chose the field of epigenetics as a case study of interdisciplinary characteristics.

II. DATA AND METHODS

A. Data

The literatures which contain the word ‘epigenetic*’ in their title, abstract or keywords were retrieved from the scientific literature database, known as ‘Web of Science’ between 1970 and 2012. Only literatures type of article and review were taken into account, removing those with the format ‘news’, ‘meetings abstract’, ‘letter’ and other non-original papers, the dataset was reduced to 30,304 records. The last date on which we retrieved data was June 7, 2013.

B. Methods

1) Co-citation Analysis

Co-citation analysis is the most influential citation analysis method, including document co-citation analysis, author co-citation and discipline co-citation analysis, etc. Here, we introduce discipline co-citation analysis, and the principles of the aforementioned analyses are similar.

The principle of discipline co-citation analysis is as follows: if two disciplines are cited together in one or more disciplines, we say that the two disciplines are co-cited. Higher co-citation frequencies indicate closer links between the disciplines. Thus, based on the discipline citation relationship, we can analyze the affiliations between disciplines. If the disciplines are divided into clusters and classes, we can analyze the interdisciplinary relationship of the research according to the contents of documents [6,7].

2) Betweenness Centrality

The betweenness centrality of a node in the network measures the importance of the position of the node in the network. It measures the extent to which the node is in the middle of a path that connects other nodes in the network. The betweenness centrality node called hub node shows where a particular discipline has connections to many different disciplines. A widely co-cited hub discipline is a good candidate for significant interdisciplinary attribute [7,8]. In this study, if the value of betweenness centrality is greater than 0.1 (>0.1), the discipline citation ring clearly shows peripheral

purple, the value of betweenness centrality is proportional to purple size in the ring.

3) Citation Bursts

A citation burst detection determines whether a given frequency function has statistically significant fluctuations during a short time interval within the overall time period. It can be also used to detect whether a particular connection has been significantly strengthened within a short period of time. In this study, if a discipline ring shows red in the middle, it represents an emerging period of a rapid development discipline [7,8].

III. RESULTS AND DISCUSSIONS

A. Interdisciplinary Span of Epigenetics

Fig.1 clearly shows a great span of epigenetics interdisciplinary and a rapid upward trend. In the 1970s, the number of interdisciplinary epigenetics is 37, in the 1980s it is 55, in the 1990s it rises to 89 and in the 21st century, it high reaches 159, which reveals apparent interdisciplinary attribute of epigenetics.

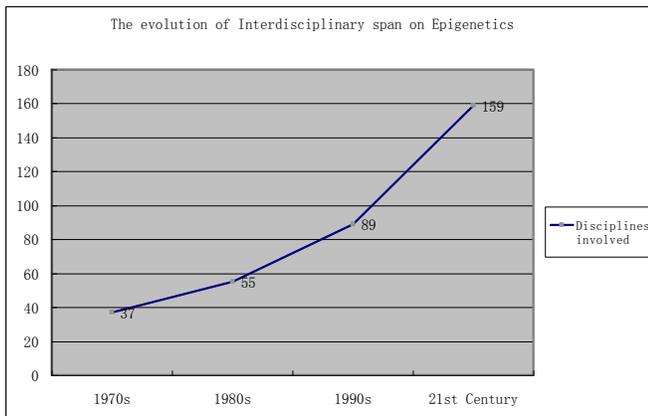


FIGURE 1. THE EVOLUTION OF INTERDISCIPLINARY SPAN ON EPIGENETICS

B. The Interdisciplinary Evolution Analysis of Epigenetics

Science knowledge maps in all steps were accomplished automatically by CiteSpace II.

Each article indexed by the Web of Science is assigned one or more subject categories. Fig.2, Fig.3, Fig.4 and Fig. 5 respectively shows a network of such subject categories, which retains the most salient connections. A citation ring represents a discipline. The size of a discipline citation ring represents the paper number of epigenetic research, the line among the different disciplines represents the overlap relationships and closeness between them.

1) The Interdisciplinary Analysis of Epigenetics in the 1970s

In the 1970s, there were 37 disciplines involved epigenetic research (Fig.1). Table 1 respectively lists top 5 disciplines of the most citations and betweenness centrality analysis, and

there are no citation bursts appearing in Fig.2. During this period, the top 5 disciplines with higher occurrence frequency were Multidisciplinary Sciences (top 1), Geochemistry & Geophysics (top2), Genetics & Heredity (top3), Cell Biology (top 4) and Zoology (top 5). The top 5 disciplines in the intermediary of interdisciplinary were Genetics & Heredity (0.11), Neurosciences (0.66), Psychology (0.03), Evolutionary Biology (0.03) and Toxicology (0.03).

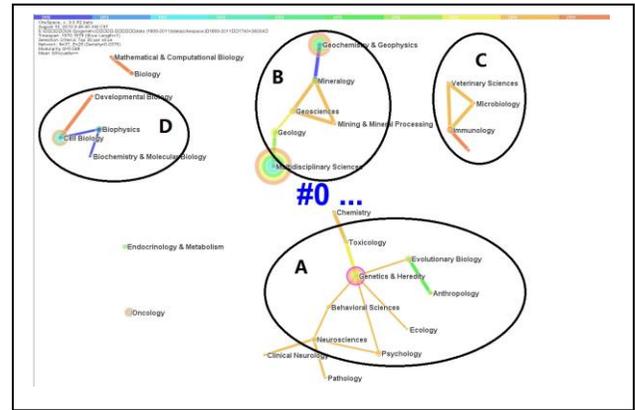


FIGURE 2. THE INTERDISCIPLINARY CONDITIONS IN EPIGENETIC RESEARCH (1970-1979)

In the 1970s, there are four discipline clusters involved in epigenetic research, i.e., A,B,C and D, which are four unconnected discipline clusters, but in a cluster internal, the disciplines are interdisciplinary. As the center of Cluster A, Genetics & Heredity shows purple in the ring with a higher value of betweenness centrality (0.11), that is, Genetics & Heredity was the hub discipline in Cluster A in this period. There were other disciplines involved in epigenetics--Chemistry, Toxicology, Evolutionary Biology, Anthropology, Behavioral Sciences, Neurosciences, Ecology, Psychology, Pathology and Clinical Neurology. Cluster B includes Geochemistry & Geophysics, Mineralogy, Geosciences, Mining & Mineral Processing, Geology and Multidisciplinary Sciences. Cluster C includes Microbiology, Veterinary Sciences and Immunology. Cluster D includes Biochemistry & Molecular Biology, Biophysics, Cell Biology and Developmental Biology. Fig. 2 also shows some small groups of subjects, such as, Biology and Mathematical & Computational Biology, they began to focus on epigenetic research respectively in 1971 and in 1979. Through analyzing our collecting data, it shows oncology began to focus on epigenetic research in 1971, but an independent one without crossing with other disciplines. In addition, Endocrinology & Metabolism was also an independent one and with its focus on epigenetics in 1974.

TABLE I. DISCIPLINES OF TOP 5 WITH THE STRONGEST BETWEENNESS CENTRALITY AND MOST CITED REFERENCES (1970-1979)

classification	Disciplines	value
mosted cited references	Multidisciplinary Sciences	16
	Geochemistry & Geophysics	10
	Genetics & Heredity	7
	Cell Biology	7
	Zoology	5
the strongest betweenness centrality	Genetics & Heredity	0.11
	Neurosciences	0.06
	Psychology	0.03
	Evolutionary Biology	0.03
	Toxicology	0.03

2) *The Interdisciplinary Analysis of Epigenetics in the 1980s*

In the 1980s, there were 55 disciplines involving epigenetic research (Fig.1). Table 2 respectively lists top 5 disciplines of the most citations and betweenness centrality analysis, and there are also no citation bursts appearing in Fig.3. During this period, the top 5 disciplines with higher occurrence frequency were Biology (top 1), Biochemistry & Molecular Biology(top 2), Cell Biology (top 3), Multidisciplinary Sciences(top 4)and Mineralogy (top 5). Fig. 3 shows there are five purple discipline rings which represents the top 5 disciplines in the hub and intermediary of interdisciplinary, i.e. Genetics & Heredity (0.33), Toxicology (0.24), Developmental Biology (0.17), Biochemistry & Molecular Biology (0.14) and Neurosciences (0.14). Compared with the ones in the 1970s, Genetics & Heredity, Toxicology and Neurosciences were still in the center of the crossing. And, Developmental Biology and Biochemistry & Molecular Biology were also the key bridge disciplines in this period.

In Fig. 3, there were eight discipline clusters, from Cluster A to Cluster H, among which, Cluster A,B,C,D and E were five disciplines crossing each other, Cluster F,G, and H were relatively independent. As one of the core disciplines, Genetics & Heredity was still in a core status in interdisciplinary. Cluster A mainly included Medicine, Toxicology, Oncology, Pharmacology & Pharmacy and Chemistry. Biochemistry & Molecular Biology grew up rapidly in this period, which was the core interdisciplinary in Cluster B, and Cell Biology, Physiology and Plant Sciences appeared in Cluster B. Through Developmental Biology, Cluster C connected with Genetics & Heredity, in this cluster, Zoology, Psychology, Neurosciences and Behavioral Sciences also began to focus on epigenetics' related researches. Cluster D includes Public, Social Sciences, Demography and Environmental Sciences, which shows humanities and social sciences began to focus on epigenetic research and illustrates in the disciplines development process, the crossing and integration of the natural science and humanities science could promote the development and improvement for the discipline itself. Some independent disciplines in the 1970s, such as, Biology and Mathematical & Computational Biology grew up rapidly in the 1980s to be interdisciplinary. Also, Cluster F composed by Geosciences,

Geology and Engineering, Cluster G and Cluster H formed independent ones involved in epigenetic in this period.

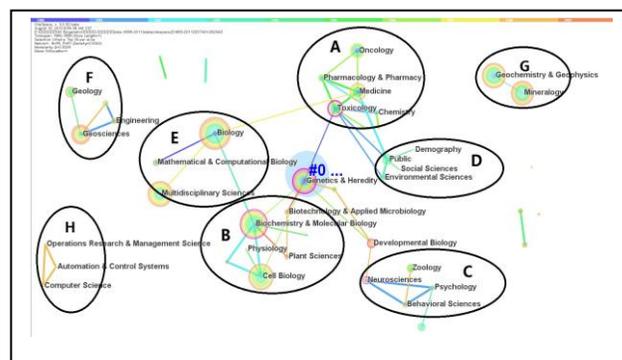


FIGURE 3. THE INTERDISCIPLINARY CONDITIONS IN EPIGENETIC RESEARCH (1980-1989)

TABLE II. DISCIPLINES OF TOP 5 WITH THE STRONGEST BETWEENNESS CENTRALITY AND MOST CITED REFERENCES (1980-1989)

classification	Disciplines	value
mosted cited references	Biology	20
	Biochemistry & Molecular Biology	17
	Cell Biology	16
	Multidisciplinary Sciences	16
	Mineralogy	15
the strongest betweenness centrality	Genetics & Heredity	0.33
	Toxicology	0.24
	Developmental Biology	0.17
	Biochemistry & Molecular Biology	0.14
	Neurosciences	0.14

3) *The Interdisciplinary Analysis of Epigenetics in the 1990s*

In the 1990s, interdisciplinary span of epigenetic increased significantly. There were 89 disciplines involving epigenetic research (Fig.1). Table 3 respectively lists top 5 disciplines of the most citations, citation burst and betweenness centrality analysis. During this period, the top 5 disciplines with higher occurrence frequency were respectively Genetics & Heredity(174), Biochemistry & Molecular Biology (167), Oncology(148), Cell Biology(131)and Geochemistry & Geophysics(104).Fig. 4 appeared citation burst discipline rings(with red in the center of the ring), the top 1 is Geosciences(4.66), which began to focus on epigenetic research in the 1970s, and developed in 1980s, and grew up to one of the strongest emerging interdisciplinary. Geochemistry & Geophysics was the top 2 (3.74).Ecology, Agronomy and Pharmacology & Pharmacy were respectively the top3,4 and 5. But because of the three discipline rings are small and they couldn't appear in Fig. 4. Fig. 4 shows there are several purple discipline rings which represents the top disciplines in the hub and intermediary of interdisciplinary, i.e. Biochemistry &

Molecular Biology (0.51), Genetics & Heredity (0.47), Developmental Biology (0.45), Biophysics (0.43) and Engineering s (0.4). Compared with the ones in the 1980s, Genetics & Heredity, Biochemistry & Molecular Biology and Developmental Biology were still the core disciplines and in the hub status. In addition, Biophysics and Engineering generated to an important cross center in this period.

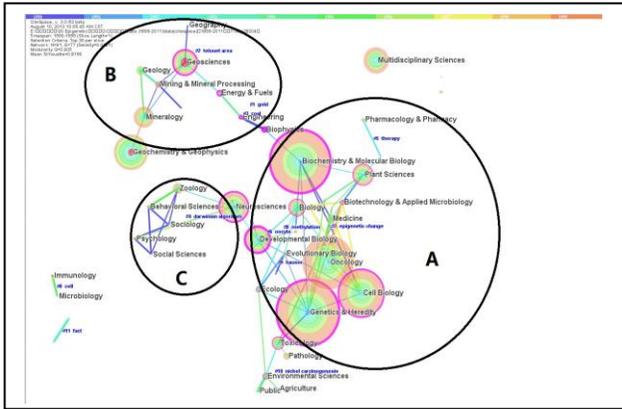


FIGURE 4. THE INTERDISCIPLINARY CONDITIONS IN EPIGENETIC RESEARCH (1990-1999)

TABLE III. DISCIPLINES OF TOP 5 WITH THE STRONGEST BETWEENNESS CENTRALITY, CITATION BURSTS AND MOST CITED REFERENCES (1990-1999)

classification	Disciplines	value
mosted cited references	Genetics & Heredity	174
	Biochemistry & Molecular Biology	167
	Oncology	148
	Cell Biology	131
	Geochemistry & Geophysics	104
the strongest betweenness centrality	Biochemistry & Molecular Biology	0.51
	Genetics & Heredity	0.47
	Developmental Biology	0.45
	Biophysics	0.43
	Engineering	0.4
citation bursts	Geochemistry & Geophysics	4.66
	Geosciences	3.74
	Ecology	3.69
	Agronomy	3.11
	Pharmacology & Pharmacy	2.91

In the 1990s, there were three main interdisciplinary clusters, i.e. A, B and C (Fig. 4), which interrelated and linked as a whole one, and compared with the ones in the 1970s and the 1980s, the three showed greater degree of association and greater cross attributes. Cluster A mainly included Biochemistry & Molecular Biology, Developmental Biology, Toxicology, Oncology and Cell Biology. It is much more worth noting that Plant Sciences and Agriculture appeared in the

cluster, which clarified the interdisciplinary fusion phenomenon of Arts, science, engineering, agriculture and medicine, etc. In Cluster B, Geosciences, Geology and Engineering, which began to focus on epigenetic in the 1980s, developed rapidly in this period, with other interdisciplinary, such as, Mineralogy, Energy & Fuels and Geochemistry & Geophysics, etc. Zoology, Psychology, Neurosciences and Behavioral Sciences made up of Cluster C. Also, Immunology and Microbiology formed a small disciplinary clusters which concerned the subject of cell.

4) The Interdisciplinary Analysis of Epigenetics in the 21st Century

In the 21st century, interdisciplinary span of epigenetic keeps increasing significantly. There were 155 disciplines involving epigenetic research (Fig.1). Fig. 5 demonstrated epigenetics in this period blended and fused into a mesh interdisciplinary, without any independent clusters, and the trends of mutual penetration and cross fusion between the disciplines was more prominent. Table 4 respectively lists top 5 disciplines of the most citations, citation burst and betweenness centrality analysis. During this period, the top 5 disciplines with higher occurrence frequency were respectively Biochemistry & Molecular Biology(4939), Oncology(4371), Cell Biology(3621), Genetics & Heredity(3472) and Medicine(1076). Figure 5 shows that the most common discipline is Biochemistry & Molecular Biology, which has the largest circle, followed by Oncology, Cell Biology, Genetics & Heredity and medicine. Multidisciplinary Sciences, with some of its citation rings in red, is a category in which the number of articles has increased rapidly, which shows that the trend of interdisciplinary study and cross-disciplinary development in epigenetics research in recent years has been more apparent. Some other emerging disciplines have too small citation rings to be clearly displayed in Figure 5.

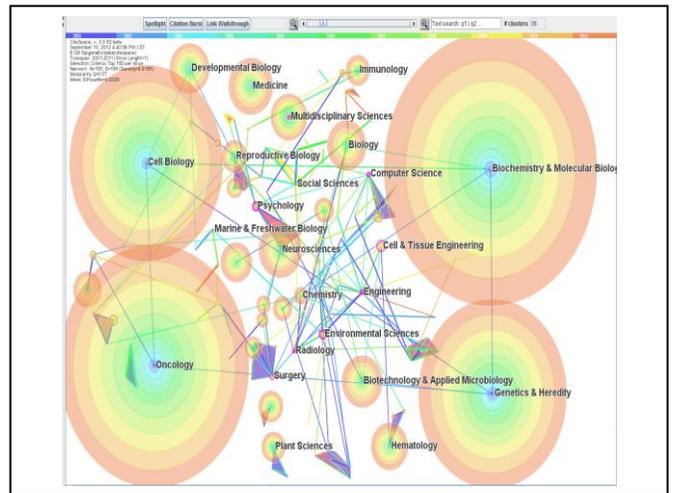


FIGURE 5. FIGURE 5 THE INTERDISCIPLINARY CONDITIONS IN EPIGENETIC RESEARCH (2000-2012)

There were also some disciplines with betweenness centrality (> 0.1) and the outer rings showed purple, such as Radiology, Computer Science, Engineering, Environmental Sciences, Psychology, Surgery, Social Sciences, and Cell &

Tissue Engineering. Although these rings are smaller themselves, the disciplines they represent play an important role in epigenetic development to be bridges and make a greater contribution to its development. Table IV respectively lists top 5 disciplines of the most citations, citation bursts and betweenness centrality analysis.

In addition, Hematology, Immunology, Reproductive Biology and Plant Sciences are relatively smaller, they are also closely related with epigenetics development.

TABLE IV. DISCIPLINES OF TOP 5 WITH THE STRONGEST BETWEENNESS CENTRALITY, CITATION BURSTS AND MOST CITED REFERENCES (2000-2012)

Classification	Disciplines	value
mosted cited references	Biochemistry & Molecular Biology	4939
	Oncology	4371
	Cell Biology	3621
	Genetics & Heredity	3472
	Medicine	1076
the strongest betweenness centrality	Radiology	0.38
	Computer Science	0.32
	Engineering	0.3
	Environmental Sciences	0.24
	Psychology	0.21
citation bursts	Multidisciplinary Sciences	25.08
	Geochemistry & Geophysics	16.09
	Geology	11.85
	Allergy	8.37
	Geosciences	6.02

IV. CONCLUSIONS

Using scientific metrology methods to analyze the features of epigenetic interdisciplinary, we found that epigenetic involved numerous interdisciplinary fields with a multi-disciplinary, multi-method and cross fusion development trend.

In the 1970s, Genetics & Heredity, Geosciences, Biochemistry & Molecular Biology and Microbiology, as core disciplines of four large discipline clusters (37 independent disciplines) and began to focus on the epigenetic research. But the four clusters are unconnected with each other. In the 1980s, the interdisciplinary features of epigenetics had come out, and been apparent in the 1990s, and in the 21st century, the interdisciplinary span of epigenetic has increased significantly. Since 1980s, in epigenetics development, with Genetics & Heredity as a leader, Biochemistry & Molecular Biology, Cell Biology and Oncology have been core disciplines, and the features of interdisciplinary has been more and more significant and the trend of multi-disciplinary has kept stable. Along with other disciplines to enter the field of epigenetic research, epigenetics gradually shows a large cross and integration development trend, beyond the core clusters there seems to have formed progressively more and more other disciplinary groups.

Epigenetics development reflects the interdisciplinary characteristics: First, natural sciences, social sciences, humanities, mathematics, science and philosophy, etc, had developed in the way of both external cross and internal cross to have formed a comprehensive and systematic body of knowledge; Second, it embodies a characteristics of interdisciplinary integration of arts, science, engineering and medicine; Third, interdisciplinary is probably to explore in depth an emerging frontier, the most significant scientific breakthroughs, which may have led to a revolutionary change in science.

Complying with the trend of cross, fusion and integration in epigenetics development, China's epigenetic researchers should learn from the interdisciplinary attribute and diversity and turn to other disciplines for more interdisciplinary theory and methods, so as to enrich and develop Chinese research direction and fields in epigenetics or other disciplines, to make epigenetics stronger and mature, and to promote various disciplines and research fields to blend and cross in a better way.

ACKNOWLEDGMENT

This work was supported by the National Nature Science Foundation of China under grant numbers 71240006, and by the soft science research projects in Shanxi Province under grant numbers 2013041082-02.

REFERENCES

- [1] Yongxiang Lu, "The significance of Interdisciplinary and cross-scientific," Chinese Academy of Science, vol.1, pp. 58-60, Jan. 2005.
- [2] Jinguang Hu and Xiong Chen, "Chinese Constitution Research Methods," Zhejiang Academic Journal, vol.4, pp. 132-137, April 2005.
- [3] Haiyan Hou, Chen C, Zeyuan Liu and Yue Chen, "The study on interdisciplinary attribute of Knowmetrics," Studies on Science in Science, vol. 98, pp. 328-332, March 2010.
- [4] Wu C T, Morris J R, "Genes, genetics and epigenetics: a correspondence", Science, vol. 293, pp.1103-1105, 2001.
- [5] Goldberg A D and Allis CD, "Bernstein E. Epigenetics: a landscape takes shape," Cell, vol.128, pp:635-638, 2007.
- [6] Zeyuan Liu, Yue Chen, Haiyan Chen, Mapping Knowledge Domains Methods and Application. 1st ed. Peking: People Publishing House, 2008, PP 34-35, PP 68-81.
- [7] Chen CM, Ibekwe-SanJuan F and Hou JH, "The structure and dynamics of cocitation clusters: a multiple-perspective cocitation analysis," J Am Soc Inf Sci Technol, vol. 61, pp:1386-1409, 2010.
- [8] Chen CM, "CiteSpace II: detecting and visualizing emerging trends and transient patterns in scientific literature," J Am Soc Inf Sci Technol, vol. 57, pp: 359-377, 2006.