Journal of Contemporary Eastern Asia, Volume 13, No.1: 19-37

Joint bibliometric analysis of patents and scholarly publications from cross-disciplinary projects: implications for development of evaluative metrics

Pitambar Gautam[1], Kota Kodama[2] and Kengo Enomoto[3]

In an attempt to develop comprehensive evidence-based methods for evaluation of the R&D performance of cross-disciplinary projects, a joint bibliometric analysis of patents and publications was performed for two industry-university-government collaborative projects aimed at commercialization: Hokkaido University Research & Business Park Project (2003-2007; 63 inventors; 176 patents; 853 papers), and Matching Program for Innovations in Future Drug Discovery and Medical Care – phase I (2006-2010; 46 inventors; 235 patents; 733 papers). Besides the simple output indicators (for five years period), and citations (from the publication date to the end of 2012), science maps based on the network analysis of words and co-authorship relations were generated to identify the prominent research themes and teams. Our joint analysis of publications and patents yields objective and mutually complementing information, which provides better insights on research and commercialization performance of the large-scale projects. Hence, such analysis has potential for use in the industry-university project's performance evaluation.

Keywords: triple helix; bibliometry; patent metrics; science mapping; cross-disciplinary; industry-university-government collaboration

Introduction

Scholarly publications and patent applications are believed to be the most easily accessible results that represent the starting objective material to judge the degree of success or failure of any large-scale university-industry-government collaborative (triple helix (TH)) research and development (R&D) project conducted at a university. While the scholarly publications as objects of a bibliometric analysis have been increasingly used in evaluating the TH projects during the interim, ex-post or follow-up evaluations together with peer reviews from experts, the patent-related data seem to be not fully explored in practice in Japan (Watanabe 2012). Anzai et al. (2012) have reviewed the current situation in Japan with regard to the evaluation of interdisciplinary research under competitive large-scale government-funded projects and the policy issues. While introducing a set of multi-dimensional key performance indicators based on research paper productivity,

[1] Creative Research Institution, Hokkaido University, Japan.
[2] Creative Research Institution, Hokkaido University, Japan; corresponding author: kota.kodama@cris.hokudai.ac.jp
[3] Patent Result,Tokyo, Japan.

they pointed to the need of using the metrics devised to evaluate (i) the patents as the technological seeds and creations of the university-led ventures, and (ii) the management activities of the large-scale and competitive TH projects. Leydesdorff (2012) and others consider the TH partnerships as neo-institutional arrangements prone to social network analysis under the TH model that can be used for policy advice about network development, for example in the case of transfer of knowledge and the incubation of new industry. According to these researchers, patents with three social coordination mechanisms (wealth generation on the market by industry, legislative control by government, and novelty production in academia) often considered as output indicators for science and technology (S&T), do actually function as input into the economy primarily to provide legal protection for intellectual property. Patents placed in this context therefore deserve proper attention both during the formulation of S&T policy as well as evaluations of R&D programs together with the scholarly publications (Nishimura and Okada 2009).

This study is a part of our ongoing program to develop new evidence-based methods for evaluation of the large-scale cross-disciplinary (a broad term that encompasses multidisciplinary, interdisciplinary, and transdisciplinary as summarized by Wagner et al. 2011) projects and exploring their potential for short to long-term S&T innovation policy formulation (Kodama and Enomoto 2014, in press). We conducted joint bibliometric analysis of patents and scholarly publications for TH projects executed at Hokkaido University (HU), a comprehensive mainstay university with dedicated educational and R&D programs. Below, we briefly introduce two projects funded mainly through the special coordination funds of the Japanese government (targeted at innovations in systemic reforms at universities and industry-university collaboration aimed at commercialization), without and with the involvement of matching funds from the partner industries (Hokkaido University 2011).

The Hokkaido University Research & Business Park Project (FY 2003-2007)

This project, referred hereafter as the HUR&BP Project in illustrations but the Business Park Project within the text, is one of the 13 projects adopted during FY2001-FY2005 by the Japanese Ministry of Education, Culture, Sports, Science and Technology (MEXT) designated for Strategic Research Centers (Super-COEs). It had the mission to materialize a system for executing activities ranging from R&D to commercialization through the combined efforts of the governing bodies of Hokkaido Prefecture, financial circles, industries, and academic institutions. Its objectives were to execute organizational innovations aimed at developing highly-competent human resources and establishing a progressive R&D system at HU, and thereby to promote knowledge creation and utilization with emphasis on utilization of the intellectual property of universities and other academic institutions in order to develop innovative technologies and new products and create business ventures. The Business Park Project promoted R&D as strategic priority programs in four interdisciplinary and integrated research areas, chosen taking into account the priorities of the 2nd Science and Technology Basic Plan of Japan (FY2001-FY2005) on life sciences, information and telecommunications, environmental sciences, and nanotechnology/materials (MEXT 2006) as well as the regional of needs of Hokkaido, as follows: (i) Development of diagnosis and treatment of zoonosis; (ii) Tissue transplantation and engineering; (iii) Sustainable and safe food production; and (iv) Policy of environment and technology. The Business Park Project was executed at the Creative Research Institution of HU (Gautam and Yanagiya 2012, Hokkaido University 2011).

The Matching Program for Innovations in Future Drug Discovery and Medical Care (Project term: FY 2006-2015, Analysis term: FY 2006-FY2010)

This project, referred hereafter as the IFDD&MC1 or the Future Drug Discovery Project, is one of the twelve projects selected under the program for “Creation of Innovation Centers for Advanced Interdisciplinary Research Areas” of the Japanese Government initiated as a part of undertakings to reinforce the system for innovation in Japan under the 3rd Science and Technology Basic Plan (FY2006-FY2011). The MEXT initiative envisaged a wide range of advanced interdisciplinary R&D, including fundamental studies relevant for future application through industry-university-government collaboration. Initially, the advanced interdisciplinary areas included engineering, medical science and pharmaceutical science, toward creating a comprehensive patient-friendly medical system as well as nano-biotechnology and information technology. Each program with a duration of ten years had three purposes: i) ensuring equal collaboration of the university and companies from the outset of program planning, ii) building the world-class research centers, in Japanese universities, dedicated to advanced interdisciplinary areas in order to achieve social and economic results in 10-15 years by developing innovative technologies and creating new industries, and iii) enhancing the international competitiveness of Japan's economy and industries. Hokkaido University proposed a program by integrating projects on (i) innovative drug discovery (Faculty of Advanced Life Science, Graduate School of Life Science, and Shionogi & Co., Ltd.) and (ii) molecular imaging and its application (Graduate School of Medicine and Hitachi, Ltd.) keeping in view a particular interest in promoting life science and the aim to achieve innovations in drug discovery and medical care focused on patient quality of life. For effective implementation, HU carried out necessary systemic reforms and established collaborative relations among the faculties of physical sciences, pharmacy and medical science toward developing a framework of education and research related to life science.

The Future Drug Discovery Project aims at developing human resources and conducting practical studies that accelerate the development of useful new drugs and technologies for diagnosis and therapies. Activities of the project are targeted at next-generation drug discovery and advanced medical care. To be specific, these include: integration of advanced biotechnology research findings and concentrating them under the program, facilitating R&D that leads to the ongoing creation of candidate next-generation drugs, and applying the results of R&D to advance medical care at the Hokkaido University Hospital. The R&D program deals with: biologically relevant compounds, such as glyco-conjugates and conjugated lipids; design and synthesis of post-translational modified proteins; structural analysis of disease-related proteins by means of NMR, X-ray imaging and photon imaging, and utilizing positron emission tomography (PET) to develop a next-generation imaging system for molecular measurement. The newly developed system is to be applied to advanced medical care, including preclinical diagnosis, functional gene/regenerative therapies and molecular target/tracking therapies, for the purpose of promoting R&D for noninvasive and patient-friendly therapies. The project was approved in 2008 after its rescreening in the third year for confirmation to its second stage. It has been executed under the direct administration of the president of HU in close cooperation from Shionogi & Co., Ltd., Hitachi, Ltd. Sumitomo Bakelite Co., Ltd., Nihon Medi-Physics Co., Ltd., and Mitsubishi Heavy Industries, Ltd.

1. Patent and publication data and methods of analysis

In view of our comparative study based on information in patent applications and journal publications, first of all, we prepared lists of target researchers who filed at least one patent and also had affiliations to one of the projects concerned. Altogether 46 researchers (inventors) were identified for the Business Park Project for 2003-2007 and 63 researchers (inventors) for the Future Drug Discovery Project for 2006-2010, respectively. The patents and publications related to these author sets were analyzed for the following metrics.

1.1 Patent metrics

A total of 176 patents applied during 2003-2007 through the Business Park Project, and 235 patents applied during 2006-2010 through the Future Drug Discovery Project (Kodama and Enomoto 2013) have been used for analysis. The patents were analyzed using Bizcruncher® software (Patent Result Co. Ltd.) to derive the basic patent-related parameters (e.g., the current state) as well as other sophisticated objective metrics such as the patent score for each patent as well as the assignee score for the each singular or collaborative assignee type.

1.1.1 Patent score

This score, calculated as PatentScore® , assumes that the patents with greater degree of attention recorded in the patent prosecution history have a higher market value and maintenance rate. The patent prosecution history incorporates several actions made by the applicant(s), Japan Patent Office or USPTO examiners, and third parties such as the competitors. There is a high degree of correlation of the patent maintenance rate with these actions including the number of views and referrals, time taken for registration (decay time), and citations, etc., which are quantified and used to estimate raw numerical scores as the measures of the degree of attention or attractiveness of patents. The raw scores are further processed to derive the standardized patent score. The patents are rated using the magnitude range of the standard score that has a value of 0-105 or even higher (for extremely attractive patents). In this study, we recognize 3 broad categories (standard patent score ranges): A (>65), B (35-65) and C (<35). The total patent score calculated by the technical field and time of filing are highly valuable for the evaluation of patents.

2.1.2 Assignee score

The assignee score is an indicator applied to judge the overall strength of a group of patents of the different types of TH members or owners (solely the university or a private company, or the possible combinations of the university, private company or other academic or governmental institution(s)). It is obtained by summing up the differential scores (i.e., the values obtained by subtracting 50 from the magnitude of the patent scores) for all patents yielding patent scores of 50 or over. Plot of the assignee score against the maximum patent score is used to further evaluate the relative merit of the owner. In such diagram, owners even with a low number of patent applications will plot in the high score region provided that there are a significant number of patents with high degree of attention (i.e., highly rated patents for which the standard patent score is over 50).

2.2 Bibliometrics related to publications and patents, and science maps

Publications from each project were acquired from the Thomson Reuters' Web of Science (WoS) SCI Expanded, SSCI and A&HCI databases using the advanced search option with string comprising "authors' affiliation to Hokkaido University" AND "authors' last names and initials joined in series by a logical OR operator" on September 3, 2013. The search was restricted to citable items comprising articles, reviews and letters in WoS core journals. The search results were exported to spreadsheets that were further scrutinized using semi-automated method to filter the records for errors arising from similarities in authors' initials. Making full use of varied information, such as authors' full names, reprint addresses, detailed affiliations and the compositions of the research teams, we were able to reduce the publications to 853 for the Business Park Project and 733 for the Future Drug Discovery Project, respectively. Bibliometric analysis of the publications was carried out using Matheo Analyzer v. 4 (a commercial decision making tool from © Matheo Software), which permits use of a dashboard, information maps, syntheses and indicators for the given set of publication data.

Science mapping for relationships among the authors in publications and inventors in the patents was performed using the symmetric network tool available in Matheo Analyzer v. 4.0 using three filters related to (i) number of items authored, (ii) number of co-occurring pairs, and (iii) connectivity defined as the number of associations. To map the relationship among the words in the titles and abstracts of publications, we used VOSviewer v. 1.5.5 (Van Eck and Waltman 2010) that works on the text corpus extracted from the WoS output file. In the case of patents with Japanese language abstracts, the text corpus in Japanese was extracted by a Text Mining (NTT Japan) tool and further mapped for the most prominent research themes in both projects.

2. Results of the analysis of patents

Basic data (i. e, the total number, state, rating based on patent score and status) on the patents as of the end of July 2012 are listed in Table 1. Additional patent indicators related to request for substantive examination, approval rate, singular or joint ownership, citations and scores are given in Table 2. While the percentage of valid patents was 59.7% for the Business Park Project with a comparatively longer (5-9 years) prosecution history, it was 74.9% for the Future Drug Discovery Project, with a shorter prosecution history (2-6 years).

Ten patents from the Business Park Project and nine patents from the Future Drug Discovery Project have been rated A implying the significant number of patents that can be judged as attractive (Fig. 1). Further differentiation by patent scores given in Table 1 places the majority of patents (86%-89%) into class B.

Plot of the assignee score versus the maximum patent score presented in Fig. 2 allows evaluation of the quality and quantity of the patents by the ownership types. For the Business Park Project (Fig. 2, left), sole ownership of either HU or a particular company has the highest number of patents with the maximum assignee score. In contrast, despite the significant number of patents with joint ownership of HU and company, both the assignee score and maximum patent score are relatively low. These observations may imply the possible problems in coordination of the industry and university partnerships, especially with regard to the projects pioneering the TH collaboration in Japan.

For the Future Drug Discovery Project (Fig. 2, right), the maximum number of patents belong to an individual company, with drastically high assignee score and moderate maximum patent score. The second largest number of patents characterizes the university-industry collaboration with the maximum patent score. In contrast with the Business Park Project, the number of patents with sole ownership of HU for this project is rather low. The high quality of patents as shown by the patent scores may imply the progress in the transfer of the rights from HU to the partner companies. This is indicative of an increase in the degree of freedom of the partner companies to commercialization. Owing to the presence of the matching funds from companies in the Future Drug Discovery Project, most of the other patents are also owned jointly by HU and companies enabling more effective collaborative R&D leading to commercialization. Overall comparison of these projects, the pioneering project and the successor that was built upon the experience of the former (Hokkaido University 2011), therefore points to the tendency of increased transfer of the partial or almost complete rights by HU to the partner companies so as to facilitate commercialization

Table 1: Basic data on patents applications by researchers from two cross-disciplinary projects.

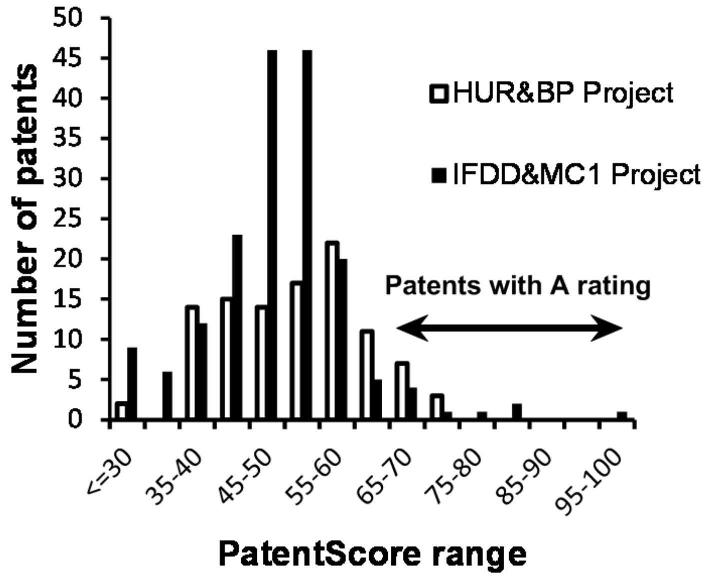
Project Name	No. of analyzed patents (5 yrs)	State of patents		Rating based on patent Score			Breakdown by status					
		valid	invalid	A	B	C	right continuation	under investigation	application Only	withdrawn	rejected	invalid
HUR&BP Project (2003-2007)	176	105	71	10	93	2	66	39	0	38	32	1
IFDD&MC1 Project (2006-2010)	235	176	59	9	152	15	56	94	26	47	11	1

Table 2: Summary of the patent indicators for 2 projects

Project Name	No. of analyzed patents (5 yrs)	**Patent indicators**									
		No. of requests for substantive examination	Ratio of requests for substantive examination	No. of patents approved	Ratio of approved patents*	No. of singular patents	No. of patents with joint ownership	Citations	Average no. of IPC per patent	Total score	Highest score
HUR&BP Project (2003-2007)	176	141	80.11%	67	65.69%	118	58	52	4.35	531.6	74.8
IFDD&MC1 Project (2006-2010)	235	167	71.06%	57	78.08%	163	72	12	3.74	545.8	97.6

Data as of 2012/7/31. *Ratio of approved patents = ($N_{\text{approved patents}}$)/ ($N_{\text{requests for substantive examination}}$ - $N_{\text{valid patents under review}}$)

Figure 1: Frequency distribution of the patent scores



Ten patents from the Business Park Project and 9 patents from the Drug Discovery Project have received an A rating defined by the PatentScore value above 65.

Figure 2: Patentee scores

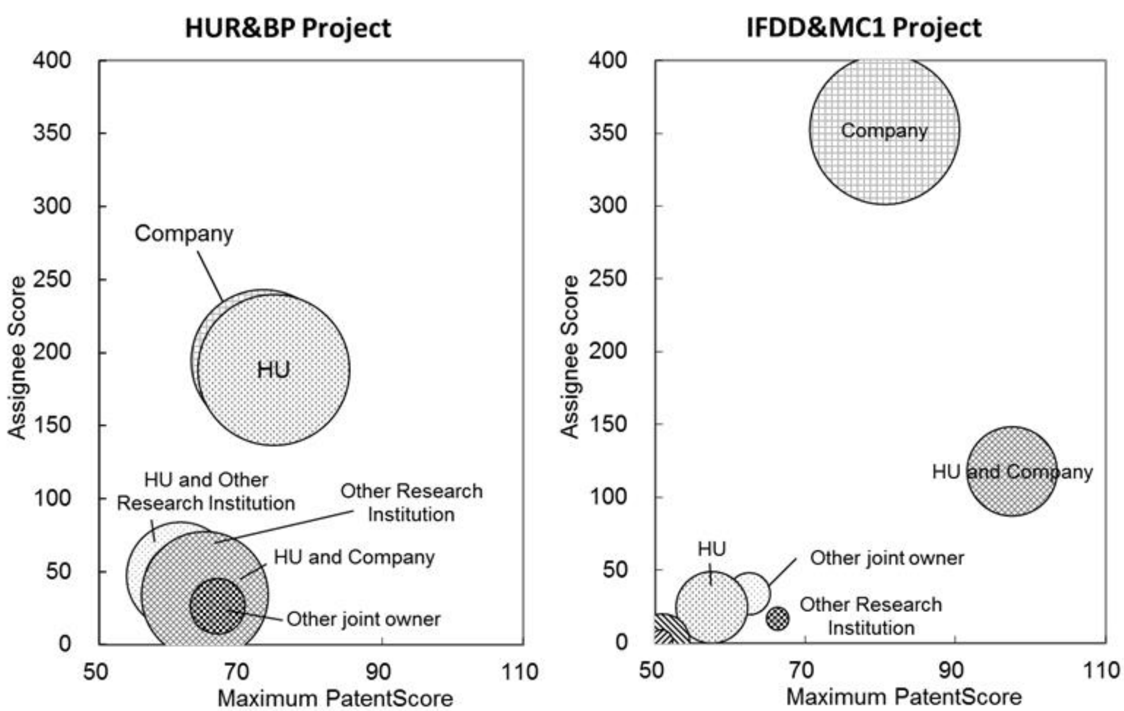


Figure 2 shows the plots of the patentee scores (sum of the differential scores (i.e. patent score – 50) of all patents yielding patent score of 60 or above) of all patents versus the calculated maximum patent score for the various assignee types for two projects. The size of the circle is proportional to the number of valid patents.

In order to demonstrate the significance the patent score as an indicator of the value of patents in terms of technology transfer, we highlight two patents with very high scores. The first patent related to the Future Drug Discovery Project and with application No. JP2007/00838, entitled "Sugar chain-capturing substance and use thereof" by Shimaoka et al. as inventors and Sumitomo Bakelite Co. Ltd. and HU as joint assignee applied on June 8, 2007, has received a patent score of 97.6. Seven citations of this patent application, such as notifications of reasons for refusal (3), notice of reasons for rejection (1), decision of refusal (2), prior-art search (1) recorded during 2011-12 in published patents and gazettes related to patents filed by Sumitomo Bakelite Co. Ltd. as assignee, contribute to its score. This patent forms the basis for a group of patents filed on the basis of the results of research on technology for manufacturing the sugar-chain capturing substances produced by the Sumitomo Bakelite Co. Ltd. including its highly popular product called the sugar chain immobilized array designed for glycosylation analysis of biologics, cells, glycans as well as the discovery of glycan biomarkers.

The second patent related to the Business Park Project has the following particulars: application No. 3855007, applied on December 2, 2003, entitled as "nanocarbon solublizer, its purification method, and high-grade nanocarbon fabrication method", invented by B. Fugetsu, assigned to HU, and with right continuation / valid status. It has received a patent score of 74.8. The reasons for such high score are 10 citations (prior-art search: 2, patent examination: 4, notices of reasons for refusal: 3, decision for rejection: 1), which appeared in published patents and patent gazettes and are related to different assignees that included individual companies, individual universities and industry-university collaborators as assignees.

Table 3: Summary of publication and citation data.

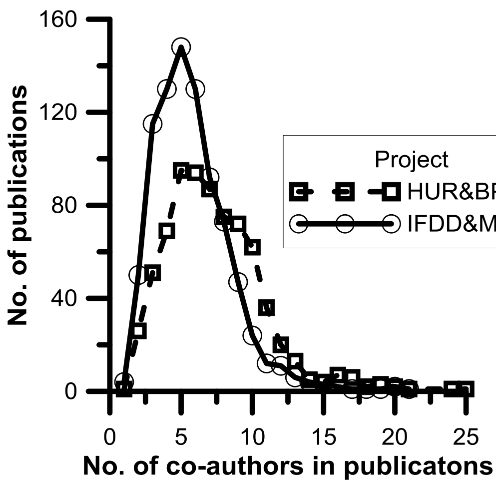
Project Name	**No. of publi-cations (5 yrs)**	**No. of co-authors per publication**			**Average annual citation per publication and share of uncited publications (until the end of 2012)**			
		min	max	average	min	max	average	% uncited
Hokkaido University Research & Business Park (HUR&BP) Project (2003-2007)	853	1	42	5.8	0	40.6	2.3	4.6
Matching Program for Innovations in Future Drug Discovery and Medical Care, Phase I (IFDD&MC1) (2006-2010)	733	1	25	7.3	0	44.4	2.9	2.0
Data on publications (Articles, Reviews, Research Notes), authorships and citations were extracted from Web of Science Database								

3. Results of basic bibliometric analysis and network analysis of scholarly publications and patents

3.1 Basic bibliometric parameters

Table 1 shows the basic data on scholarly publications in WoS core journals. There are obvious differences in the number of co-authors in the two datasets: the project with focus on drug discovery and medical care and hence entirely the biomedical field has a higher average number of co-authors per publication than the Business Park Project that covered multiple themes including nanotechnology, chemistry, physics, biomedicine, environmental public policy, etc. (see also the most frequently occurring keywords shown in Table 4 later). The difference in co-authorship pattern is clearly illustrated in Fig. 1. Also, relatively higher average citations and better citedness characterize the Drug Discovery Project, although the differences in disciplinarity and the passage of time do not permit true comparison of these projects.

Figure 3: Co-authorship patterns for two projects.



3.2 Results of network analysis and science mapping

In this section, we first describe the results of analysis of the text corpus in the titles and abstracts of the publications with interpretation in terms of the prominent research themes. Then we proceed with the analysis of co-authorships to extract information to distinguish the most productive collaborative research team compositions. Similar analysis is extended to the patent data to gain insight on the fields reflected in the summary of the patents and the co-inventor networks.

3.2.1 Results of analysis of the Business Park Project

3.2.1.1 Map of words and noun phrases in scholarly publications

Figure 4 is a science map showing the clusters defined by words and noun phrases in the text corpus formed by the words in titles and abstracts of publications. Each word and noun phrase is counted once. The diagram shows the relationship among words and noun phrases that occur at least eight times. The research themes covered by six prominent clusters (anti-clockwise from the lower left part) deciphered from the most frequently occurring words (accompanied with the weight assigned) carrying the information are as follows:

(i) Zoonotic infections and dermatitis: gene (97), patient (85), virus (59), mutation (49), strain (41), skin (40), antibody (39), infection (36), influenza (28), epidermolysis bullosa (23) , keratinocyte (22)

(ii) Protein structure analysis: protein (116), disease (97), mouse (43), receptor (42), mrna (37), brain (32), neuron (26), inhibition (23), stress (23), kinase (21), phosphorylation (17)
(iii) Plant rhizosphere: plant (36), root (26), leafe (19), phosphorus (19), soil (17), performance (16), rice (16), nitrogen (15), organic matter (9), shoot (9)
(iv) Tissue engineering: structure (129), surface (92), molecule (63), substrate (49), film (47), crystal (42), spectroscopy (42), fabrication(20), honeycomb(14), polymer solution, self-organization (9)
(v) Synthesis of polymers including gels: synthesis (55), polymer (40), particle (36), light (23), solvent (23), gel (16), size exclusion chromatography (9), deacetylation (8), phenyl (8), static laser light (8)
(vi) Regenerative medicine: animal (31), fibroblast (20), reconstruction (19), anterior cruciate ligament (18), graft (18), tendon(18), knee(17), rabbit (16), surgery (16), patellar (14), mechanical property (12)

Cluster (i) is composite of two separate research themes dealing with zoonosis and dermatology. Decreasing the threshold for the occurrence of words enables separation of a larger number of clusters corresponding to smaller research themes. In addition, joint use of the zooming option in VOSviewer helps in recognizing the details of the smaller research themes, though visualization in a single map has limitations. The research theme defined by cluster (ii) related to the protein structure analysis and neuroscience was pursued as a part of the research on genome dynamics and neuroscience research groups, some members of which had parallel appointment in the Business Park Project. Therefore, this complex theme is not directly related to the four strategic themes promoted by the Business Park Project.

Figure 4: Label map with clusters

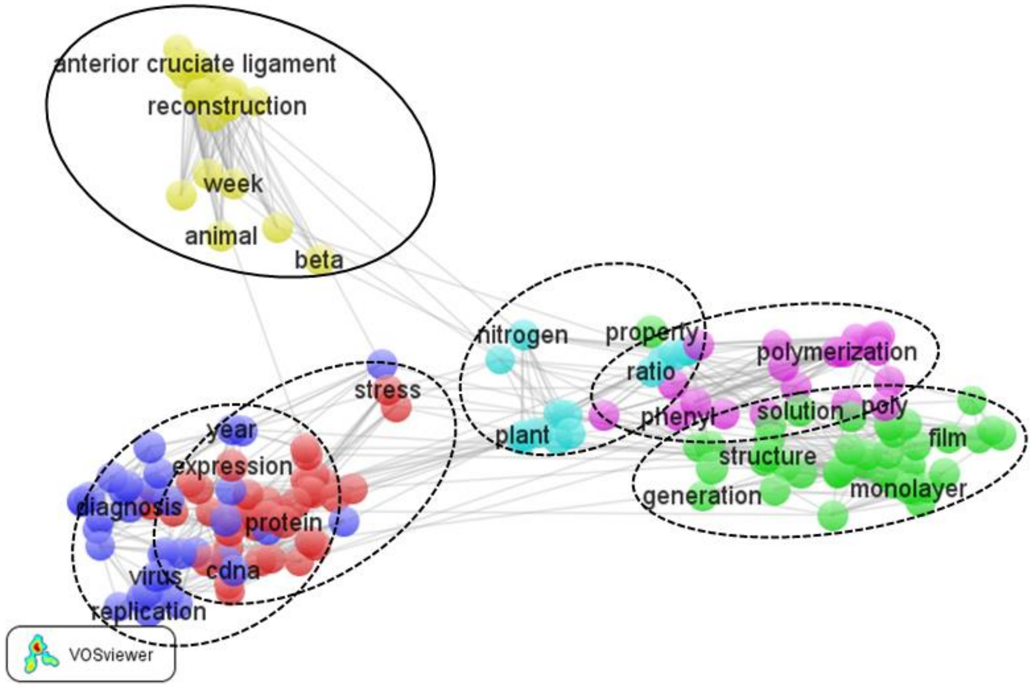


Figure 4 shows the corresponding label map with clusters (marked by ellipses) of words and noun phrases occurring more than eight times in the text corpus in titles and abstracts in 853 publications from the Business Park Project.

3.2.1.2 Co-authorship relations in scholarly publications

Figure 5 is a schematic grouping of the strongly associated networks of co-authors (minimum number of occurrence, shown in the right side, of individual author as a node = 9, minimum number of the pairs of co-authors connected or edges connected by lines = 5) in publications. Seven major clusters are evident, and the central one comprises of a strong sub cluster seen in the lower right side. These clusters can actually be interpreted as the research teams led by T. Satoh (polymer synthesis), H. Matsui (functional foods), T. Shinano (rhizosphere), H. Shimizu (dermatitis) with a sub-cluster led by K. Yasuda (bone regeneration), T. Naito (nano-devices), Y. Nomura (protein functions) and H. Kida (zoonotic infections).

Figure 5: Co-author relationships

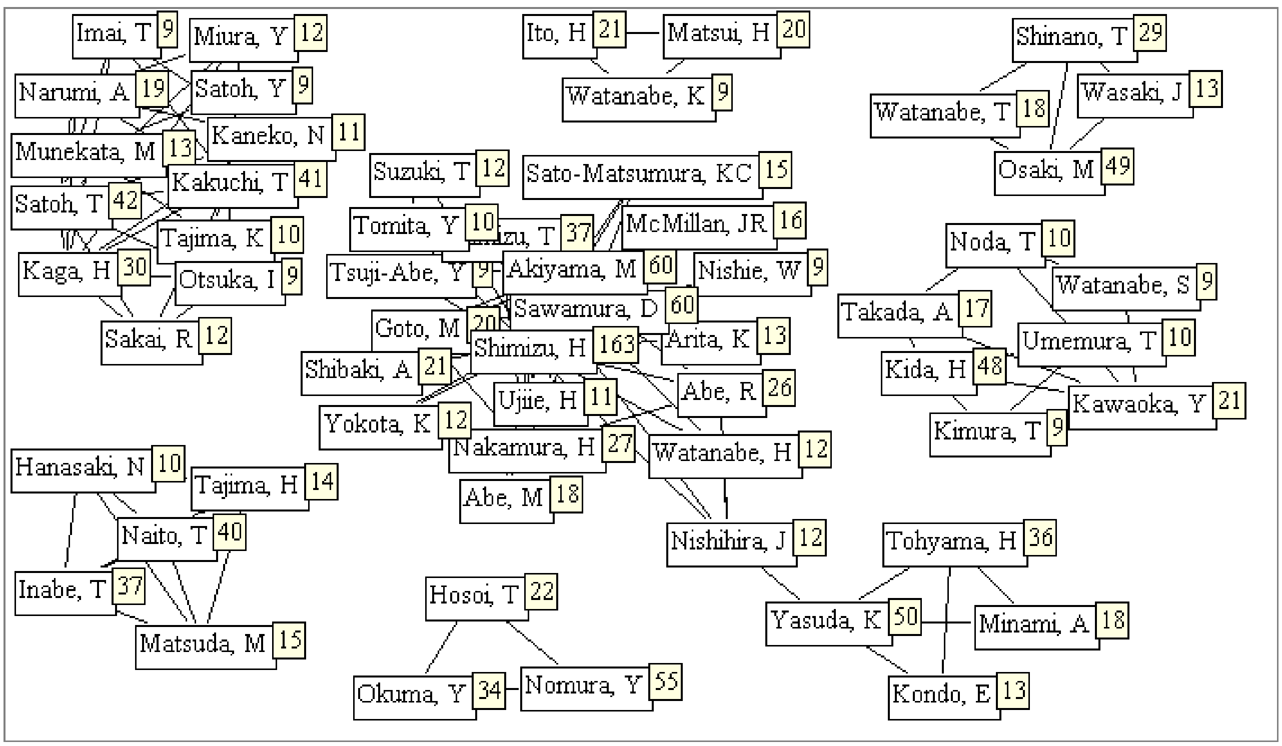


Fig. 5 Co-author relationships (minimum author threshold = 9, minimum co-authorship pair = 5) in publications from the Business Park Project. Map generated using Matheo Analyzer v. 4.0.

3.2.1.3 Text mining of the Japanese words in patent abstracts

Analysis of the Japanese words in the patent abstracts revealed 3 major strongly connected networks for the Business Park Project related to 4 research themes: carbon nanotubes, functional foods, bone-regeneration material, and diagnosis of infectious diseases. Research on carbon nanotubes was pursued as a part of innovative research programs rather than as a part of the Business Park Project, at the Creative Research Initiative Sousei that was the testbed of the project. This topic is far better represented in patent applications than in scholarly publications. In contrast, research on functional foods was a part of the Business Park Project's "sustainable and safe food production" theme that also included research on plant rhizosphere. The theme functional foods appears as a prominent field in the patent applications only and it can be explained from its high degree of relevance to commercialization.

3.2.1.4 Co-inventor relationships in patents

Figure 6 shows six distinct clusters seen as the networks of co-inventors, each having at least 5 patents and collaboration with at least 3 others, irrespective of the affiliation of the collaborator. For the sake of visibility, inventors affiliated to the studied project are identified by names, whereas others, irrespective of their affiliation to HU or an external institution or a company, are shown as X followed by a number that is unique to an individual.

The most prominent network in the map (Fig. 6, right half), with 11 inventors dealing with the fusion of bio and nanotechnology for medical applications, can be divided into 3 sub-clusters of varying sizes: (i) tissue engineering: Shimomura, Tanaka, Yamamoto, Ijiro and others; (ii) diagnosis and prevention of the zoonotic diseases: Mukasa, Takeda, Hosoi, Sawamura, Kida and others; (iii) functional foods: Sashima and others. The next two author networks, that are intimately related to the network just described but appear as distinct owing to the rather high threshold for inventors, deal with dermatitis (Shimizu and others) and polymers as biomedical materials (Satoh and others). Remaining 3 networks can be related to plant rhizosphere (Matsui, Osaki and others), dispersion of carbon nanotubes and their applications (Fugetsu and others), and biogels and cartilage regeneration (Yasuda and others).

Figure 6: Co-inventor relationships

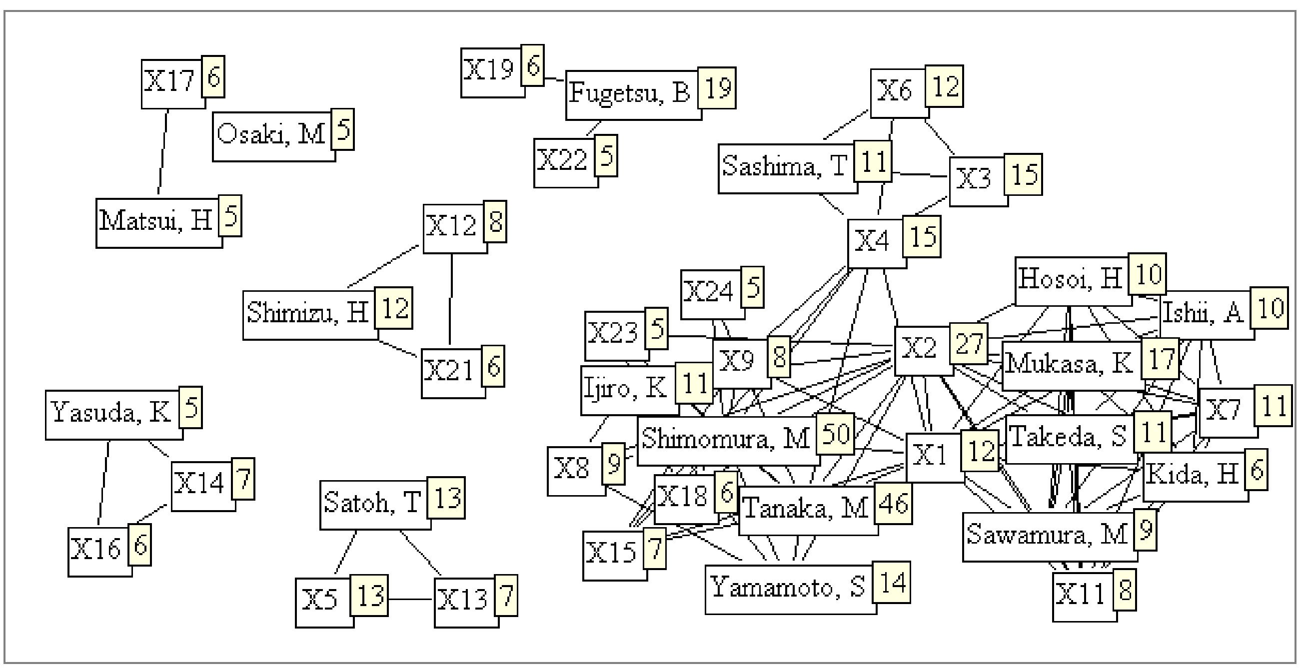


Fig. 6 Co-inventor relationships (minimum no. of occurrence of individual inventor, shown to the right end of the label = 5, minimum no. of co-inventor pairs = 3) in patents from the Business Park Project.

Figure 7: Label map with clusters

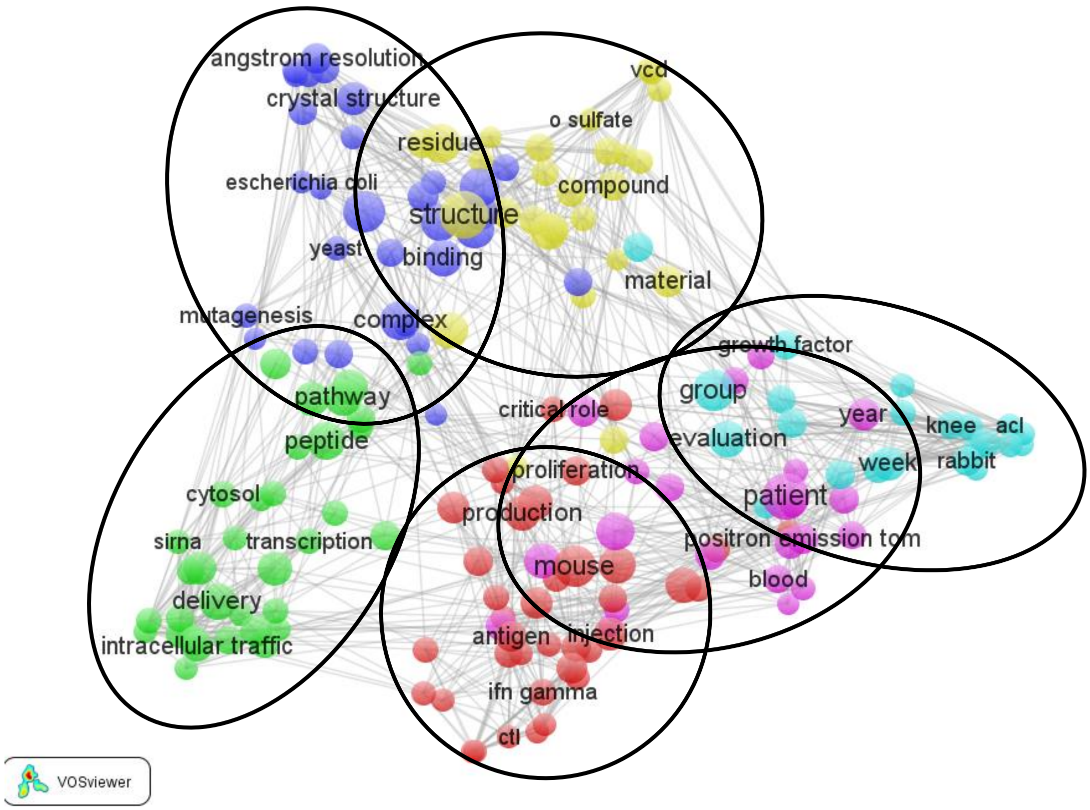


Fig. 7 Label map with clusters (marked by ellipses) of words and noun phrases in the text corpus formed by the words in titles and abstracts in 733 publications from the Future Drug Discovery Project. Map generated by VOSviewer v. 1.5.5.

3.2.2 Results of analysis of the Future Drug Discovery Project

3.2.2.1 Analysis of words and noun phrases in scholarly publications

A VOSviewer label map with clusters formed by words and noun phrases occurring at least 10 times in the text corpus of the publication titles and abstracts is presented in Fig. 7. The research themes covered by six clusters are identified (anti-clockwise from the lower left of the diagram) below based on the selected set of frequently occurring words (with assigned weights) that carry subject-specific information.

(i) Drug delivery: peptide (55), delivery (52), plasmid dna (35), liposome (32), intracellular trafficking (26), gene expression (24), fluorescence correlation spectroscopy (20), cytosol (17), green fluorescent protein (16), adenovirus (11)

(ii) Tumor diagnosis: mouse (76), tumor (50), phosphate (30), antigen (29), t cell (26), cytokine (21), rat (19), sphingosine (19), cytotoxicity (16), interleukin (15)
(iii) PET applications in therapeutic care: patient (112), imaging (43), positron emission tomography (27), survival (25), age (23), diagnosis (22), blood (20), chemotherapy (16), tomography (14), therapeutic effect (12)
(iv) Regenerative medicine: growth factor (24), rabbit (23), reconstruction (23), knee (22), surgery (19), fibroblast (15), graft(15), animal (14), angiogenesis (11), anterior cruciate ligament (11),
(v) Glycomics: structure (122), glycosylation(26), n glycan (25), oligosaccharide (25), glycopeptide (22), proteoglycan(14), vibrational circular dichroism (14), glycosaminoglycan (12), flight mass spectrometry (10)
(vi) X-ray crystal structure analysis: domain (75), enzyme (59), binding (52), crystal structure (37), mutant (22), autophagy (21), rana (20), cytoplasm (16), saccharomyces cerevisiae (13), escherichia coli (11)

3.2.2.2 Co-authorship analysis in scholarly publications

Prominent networks of prolific authors (minimum occurrences for individual and pairs being 12 and 5, respectively) in publications is shown in Fig. 8. These networks are represented by at least 1 member from the Future Drug Discovery Project: Inagaki (Nuclear magnetic resonance spectroscopy), S.I. Nishimura (glycomics), Harashima (drug delivery), Kondo (pharmaceutics) including a sub-cluster led by Yasuda (bone regeneration), Tamaki with Kuge and Katoh (radioisotope molecular imaging), and I. Tanaka with T. Nishimura (immunology).

Figure 8: Networks of co-authors

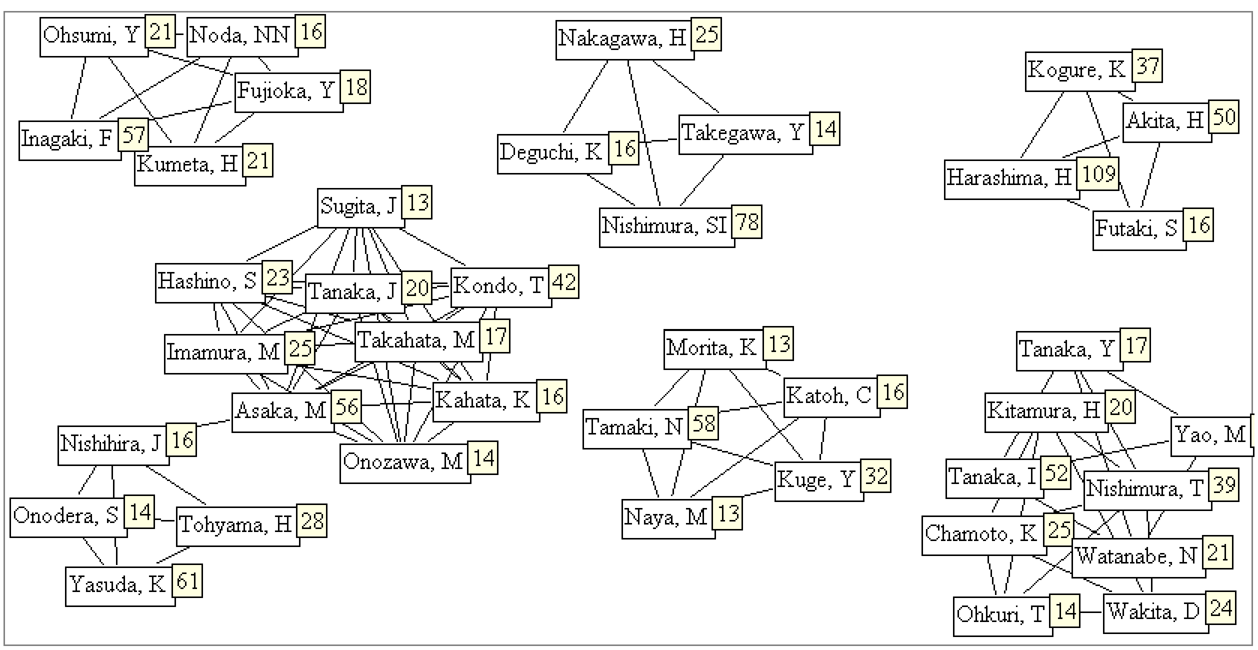


Fig. 8 Networks of co-authors, each occurring at least 12 times and each pair occurring 5 times or more), in publications from the Future Drug Discovery Project represent the strongest research teams.

3.2.2.3 Text mining of the Japanese words in patent abstracts

Network analysis of the Japanese words in the patent abstracts revealed 3 major themes: radiotherapy, RI probe with low molecular compounds, and glycomics in the order of decreasing importance.

3.2.2.4 Co-inventor relationships in patents

Figure 9 shows major clusters mapped through relationship of patent co-inventors. Several clusters, however, are represented by singular inventors and this is an artefact of the thresholds set for the individual inventors and their pairs during mapping to achieve clarity in visualization. In fact, the research teams in each cluster have several members, which can be recovered by lowering the threshold leading to a merger of smaller clusters to a large-sized cluster. A large cluster related to glycomics led by S.I. Nishimura in the lower left region in Fig. 9 is closely related to the smaller cluster led by Fujiwara situated in the lower left corner. Three moderate to big clusters in the right half of the map are led by Ishitsu (detective semiconductors), Bito (MRI imaging) and Kuge (with Tamaki; radioisotope molecular imaging). A prominent cluster led by Harashima and 2 smaller clusters related to Yasuda (regenerative medicine) and T. Nishimura (immunology), respectively, are evident as was the case in the scholarly publications (Fig. 8) although at differing scales.

Figure 9: Co-inventor relationships

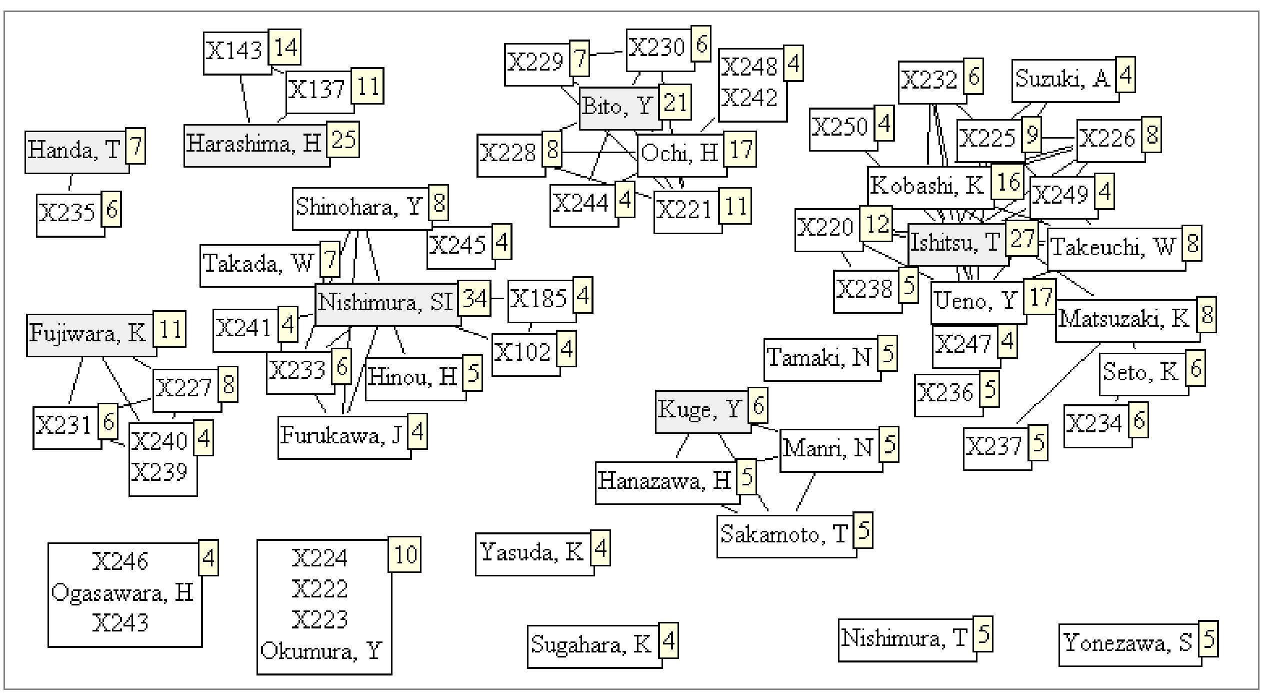


Fig. 9 Co-inventor relationships (for 4 or more individuals and 3 or more pairs of inventors) in patents from the Future Drug Discovery Project.

4. Overall interpretation and discussion

Interpretation of the multiple pieces of information derived from patents and the scholarly publications requires an integrated approach with due consideration to the several facts: (i) the majority of researchers had primary affiliation in one of the departments or research institutes of HU beside the additional fixed-term appointment in one of the projects studied; (ii) Some of the researchers with significant research output from their primary affiliation had merely administrative roles in these projects, especially with regard to the Business Park Project; and (iii) there were no strict requirements for authors to provide the project level descriptions in the patents or publications, implying that strictly assigning the publications to the particular project in question is practically difficult. As a result, the patent applications and particularly the scholarly publications may represent the total contributions by researchers having parallel appointments at HU.

Table 4: Major keywords in publications and the prolific authors in publications and patent applications

	Most frequent keywords in publications		Prolific authors in publications		Prolific inventors in patents	
S.No.	**HUR&BP**	**IFDD&MC1**	**HUR&BP**	**IFDD&MC1**	**HUR&BP**	**IFDD&MC1**
1	Chemistry (163)	Biochemistry & Molecular Biology (249)	Shimizu, H (163)	Harashima, H (109)	Shimomura, M (50)	Nishimura, SI (34)
2	Dermatology (124)	Chemistry (113)	Uosaki, K (57)	Nishimura, SI (78)	Tanaka, M (46)	Ishitsu, T (27)
3	Physics (94)	Pharmacology & Pharmacy (97)	Nomura, Y (55)	Yasuda, K (61)	Fugetsu, B (19)	Harashima, H (25)
4	Materials Science (70)	Biophysics (70)	Yasuda, K (50)	Tamaki, N (58)	Mukasa, K (17)	Bito, Y (21)
5	Biochemistry & Molecular Biology (69)	Cell Biology (53)	Osaki, M (49)	Inagaki, F (57)	Yamamoto, S (14)	Ueno, Y (17)
6	Polymer Science (67)	Radiology, Nuclear Medicine & Medical Imaging (50)	Kida, H (48)	Igarashi, Y (54)	Satoh, T (13)	Ochi, H (17)
7	Engineering (65)	Oncology (46)	Shimomura, M (48)	Tanaka, I (52)	Shimizu, H (12)	Kobashi, K (16)
8	Biotechnology & Applied Microbiology (42)	Immunology (36)	Satoh, T (42)	Kinjo, M (48)	Ijiro, K (11)	Fujiwara, K (11)
9	Neurosciences & Neurology (38)	Orthopedics (35)	Naito, T (40)	Kondo, T (42)	Sashima, T (11)	Okumura, Y (10)
10	Plant Sciences (35)	Engineering (32)	Ito, E (35)	Nishimura, T (39)	Takeda, S (11)	Takeuchi, W (8)
11	Orthopedics (34)	Hematology (29)	Watanabe, Y (32)	Sugahara, K (38)	Hosoi, H (10)	Shinohara, Y (8)
12	Virology (33)	Genetics & Heredity (28)	Shinano, T (29)	Kuge, Y (32)	Ishii, A (10)	Matsuzaki, K (8)
13	Food Science & Technology (31)	Crystallography (23)	Matsui, H (20)	Monde, K (27)	Sawamura, M (9)	Takada, W (7)
14	Environmental Sciences & Ecology (29)	Surgery (23)	Ito, S (20)	Takeda, H (25)	Kida, H (6)	Handa, T (7)
15	Agriculture (27)	Sport Sciences (22)	Tanaka, M (19)	Wakita, D (24)	Matsui, H (5)	Seto, K (6)
16	Water Resources (26)	Biotechnology & Applied Microbiology (21)	Tanaka, J (16)	Yamada, S (23)	Osaki, M (5)	Kuge, Y (6)
17	Pharmacology & Pharmacy (25)	Materials Science (21)	Kimura, K (15)	Hinou, H (20)	Yasuda, K (5)	Yonezawa, S (5)
18	Sport Sciences (24)	Research & Experimental Medicine (20)	Toda, Y (14)	Shinohara, Y (20)	Hattori, S (4)	Tamaki, N (5)

Note: Numbers in brackets indicate the no. of occurrences of keywords and authors/inventors affiliated to the Projects.

Comprehensive interpretation of the patents and publications on datasets can be made by taking into account additional data extracted from bibliometric analysis. One possible dataset is related to the keywords and prolific authors in patents and publications shown in Table 4. Keyword data demonstrate the degree of disciplinarity: the Business Park Project has a broader coverage of subjects such as chemistry, physics, engineering, plant science, agriculture, environmental science and various medical fields, while the Future Drug discovery Project has focus on biochemistry & molecular biology, medical care especially with focus on nuclear imaging and radiotherapy, immunology, etc. Clear differences are seen in authorships among the scholarly publications and patents: the most prolific authors in the publications are not necessarily the most prolific inventors. Authors with managerial roles in the projects are seen as prolific in scholarly publications but not in the patents. In contrast, some prolific inventors in patents are absent in scholarly publications as governed by the differences in the publication practices among the various TH participants. Furthermore, if the results of analysis are used to evaluate or benchmark the performance of the particular group of researchers within a project or several projects, it is recommended to calculate other bibliometric parameters, e.g., the field weighted citation impact that corrects for the discrepancies in citation practices in different disciplines, in addition to the scholarly output volume and other indicators that represent the power metrics.

Table 5: Summary of the joint analysis of the patents and publication data

Target of analysis	(1) Research & Business Park Project (2003-2007)	(2) Future Drug Discovery Project (2006-2010）	Remarks
No. of researchers/inventors (Nr/i), JPO patents (Npat) and scholarly publications in WoS core journals (Npubl)	N(r/i) = 63; N(pat) = 176; N(publ) = 853	Nr/i = 46; Npat = 235; Npubl = 733	High score patents in both but (2) has some extremely attractive patents; Higher average citation per paper in (2) due to focus on clinical medicine
Words and noun phrases networks in titles and abstracts of publications as indicators of research themes	Zoonoses and dermatitis; Protein structure; Rhizosphere; Tissue engineering; Polymer/gel synthesis; Regenerative medicine	Drug delivery; Tumor diagnosis; PET applications in therapeutic care; Regenerative medicine; Glycomics; Crystal structure	Publications highlight major research themes (from words) and teams (from author networks) but the specialized developmental topics and teams linked to commercialization may remain hidden. In contrast, text mining of the patent titles and abstracts and mapping inventor relationships gives exact information related to commercialization. For project (1), functional foods and carbon nanotubes that are distinct in patents are not apparent in the scholarly publications.
Mining the Japanese text in patents as indicators of topics related to commercial application	Carbon nanotubes; Functional foods; Bone-regeneration material; Diagnosis of zoonotic infections	Radiotherapy; RI Probe & low molecular compounds; Glycomics	
Author networks in publications as reflections of major research teams	Polymer synthesis; Functional foods; Rhizosphere; Dermatitis and bone regeneration; Nano-devices; Zoonotic infections.	NMR spectroscopy; Pharmaceutics and bone regeneration; Radioisotope molecular imaging; immunology	
Inventor networks in patents as research teams specialized in commercial applications	Tissue engineering; Zoonoses;Functional foods; Dermatitis, Polymers; Diagnosis of Influenza; Rhizosphere; Carbon nanotubes; Biogels and cartilage regenertion	Detective semi-conductors; MRI imaging; RI Probe and molecular imaging; Drug delivery; Regenerative medicine; Glycomics	

In the case of patent applications, comparison of the Business Park Project and the Future Drug Discovery Project reveals a tendency of an increased shift of ownership rights mainly by the university during the former project as the pioneering one to the joint ownership with industries during the latter project that in fact is a successor utilizing the experience gained by the former project. Such a tendency is in line with the university's policy of more effectively facilitating commercialization.

Table 5 summarizes the interpretation based on the authorship and word relations. While the authorship maps reflect major research teams, the word and noun phrase maps provide information on the major research themes. Seemingly single clusters sometimes are composite research themes (such as cluster (i) combining zoonosis and dermatitis in the case of the Business Park Project). Research theme of the regenerative medicine related to bone and ligament reconstruction is prominent in scholarly publications in both projects as well as the authorship maps derived from patent applications as the project leader pursuing this theme (K Yasuda, one of the most prolific authors) is the participant in both projects. Also, the clusters may partially overlap, as in the case of cluster (ii) & (iii), and (iii) & (iv) in the case of the Business Park Project. However, use of the zooming capacity available in VOSviewer enables visualization of minute details regarding the extent of each cluster and appreciation of the degree of overlap with the adjacent clusters.

5. Conclusions

Joint analysis of the patent applications and the scholarly publications in core journals for two cross-disciplinary university-led TH projects leads to the following observations:

(i) Patent metrics, especially the patent score. are useful in evaluating the attractiveness of the applied patents and categorizing them. The maximum patent score and the assignee score for a group of patents belonging to a single TH partner, joint or triple partnerships can be used to establish general qualitative aspects and trends in ownership relations.

(ii) Bibliometric indicators based on publications can be used to judge the research output and the citation-based impact of the projects. Science maps based on the relationship of words in titles and abstracts reveal the prominent research themes. In contrast, similar maps constructed from the co-authorship relations give clues to the composition of the research teams.

(iii) Compared to the core journal publications, patent data reveal more granular structure comprised by clusters of relatively few researchers/inventors. In R&D projects targeted at commercialization, scholarly publications in core journals are more likely to report the findings of fundamental research whereas any material related to development and practical application will be first used in applying patents. Hence, only a joint analysis of the patents and publications will yield a true picture of the research outcomes of the TH projects especially during the period of execution or soon after the completion of 3-5 year long projects.

These findings reinforce the importance of conducting joint bibliometric analysis and science mapping of the patents and publications data from any large-scale collaborative industry-university-government project, including the results of such analysis in reports prepared by those executing the project, and actively utilizing them during interim, ex-post or the follow-up assessments by the project evaluators. A wide range of indicators, such as the patent score, assignee score, share of publications in top percentiles (1%, 10% etc.) by citations or journals, field-weighted citation indices can be used simultaneously besides the output volume for the comprehensive evaluation of individual projects and for benchmarking a group of projects that fall into the same disciplinary category or funding scheme. As the objective evidence-based information, the bibliometric and patent-based indicators have potential to serve as a part of the evaluative metrics.

References

Anzai Tomohiro, Ryoichi Kusama, Hiroyuki Kodama, and Shintaro Sengoku (2012), Holistic observation and monitoring of the impact of interdisciplinary academic research projects: An empirical assessment in Japan, in: Technovation 32, 345-57.

Gautam Pitambar, and Ryuichi Yanagiya (2012). Reflection of cross-disciplinary research at Creative Research Institution (Hokkaido University) in the Web of Science database: appraisal and visualization using bibliometry, in: Scientometrics, 93(1), 101-111.

Hokkaido University (2011), Hokkaidou daigaku souseikenkyuukikou 9-nenkan no ayumi [Nine years of Creative Research Institution (CRIS)], p. 125+, Sapporo: Hokkaido University (in Japanese)

Kodama Kota, and Kengo Enomoto (2014), Development of novel evaluation method of university-industry collaboration project, in: Aspects of Intellectual Property Innovation Research, in press (in Japanese).

Leydesdorff Loet (2012), The Triple Helix of University-Industry-Government Relations, (February 2012)), online: http://eprints.rclis.org/16559/1/The%20Triple%20Helix%20of%20University-Industry-Government%20Relations.Jan12.pdf#search='The+Triple+Helix+of+UniversityIndustryGovernment+Relations' (February 2, 2014).

MEXT (2006). White Paper on Science and Technology 2006, online: http://www.mext.go.jp/english/whitepaper/1302722.htm (February 2, 2014).

Nishimura Junichi, and Yosuke Okada (2009), University-Industry-Government Collaboration in Bio-Clusters : Policy evaluation based on patent data, in: The Journal of Science Policy and Research Management, 24(4), 383-399 (in Japanese).

Van Eck Nees Jan, and Ludo Waltman (2010), Software survey: VOSviewer, a computer program for bibliometric mapping, in: Scientometrics, 84(2), 523-538, program online: http://www.vosviewer.com/ (February 2, 2014).

Wagner Caroline S., David J. Roessner, Kamau Bobb, Julie Thompson Klein, Kevin W. Boyack, Joann Keyton, Ismael Rafols, Katy Boerner (2011), Approaches to understanding and measuring interdisciplinary scientific research (IDR): A review of the literature, in: Journal of Informetrics, 5(1),14-26.

Watanabe Toshiya (2012), University Industry Cooperation Management Based On Field Experience, in: Journal of the Japan Society for Intellectual Production 8(2), 24-30 (in Japanese).