



2013

SCIENTOMETRICS

**14th International
Society of Scientometrics
and Informetrics Conference**
15th - 19th July 2013
Vienna, Austria

PROCEEDINGS Volume I

**PROCEEDINGS OF
ISSI 2013
Vienna**

VOLUME 1

14th International Society of
Scientometrics and Informetrics Conference

Vienna, Austria
15th to 20th July 2013

Editors

Juan Gorraiz, Edgar Schiebel, Christian Gumpenberger, Marianne Hörlesberger,
Henk Moed

Sponsors

ASIS&T, USA

Elsevier B.V.

EBSCO Information Services, USA

Federal Ministry for Science and Research, Austria

Federal Ministry for Transport, Innovation and Technology, Austria

Information Assistant, Verein für Informationsmanagement, Vienna

ORCID, Inc.

Science-Metrix/R&D Reports

Swets Information Services

Thomson Reuters

ZSI - Centre for Social Innovation, Vienna

All rights reserved.

© AIT Austrian Institute of Technology GmbH Vienna 2013

Printed by Facultas Verlags- und Buchhandels AG,
Stolbergasse 26, A-1050 Wien

ISBN: 978-3-200-03135-7

ISSN: 2175-1935

KNOWLEDGE CAPTURE MECHANISMS IN BIOVENTURE CORPORATIONS: A CASE STUDY

Thomas Gurney¹, Antoine Schoen², Edwin Horlings³, Koichi Sumikura⁴, Patricia Laurens⁵, Peter van den Besselaar⁶, and Daniel Pardo⁷

¹ *t.gurney@rathenau.nl*, ³ *e.horlings@rathenau.nl*

Rathenau Institute, Science System Assessment, Anna van Saksenlaan 51, 2593 HW, The Hague (The Netherlands)

² *a.schoen@esiee.fr*, ⁵ *laurens@esiee.fr*

Université Paris-Est, ESIEE – LATTIS – IFRIS, 2, bd Blaise Pascal, Noisy le Grand, 93160 (France)

⁴ *sumikura@grips.ac.jp*

GRIPS - National Graduate Institute for Policy Studies, 7-22-1 Roppongi, Minato-ku, Tokyo 106-867, (Japan)

⁶ *p.a.a.vanden.besselaar@vu.nl*

VU University Amsterdam, Network Institute & Department of Organization Science, De Boelelaan 1105, Amsterdam (The Netherlands)

⁷ *daniel.pardo@wanadoo.fr*

CNRS - Aix-Marseille Université, LEST UMR 7317, 35 avenue Jules Ferry, 13626 Aix en Provence Cedex 01 (France)

Abstract

Mechanisms of knowledge transfer from academia to industry have long been debated. The knowledge inputs required may stem from research conducted many years prior to a technology being adopted and adapted by industry, and a supporting base of knowledge is required to facilitate this. In this case study we utilise the publishing and patenting history of an individual scientist, and link their output to the technologies with which the scientist is involved. A detailed description of knowledge sources of these technologies is discussed, including the role absorptive capacity plays in priming their development. This study addresses the contributions of the researcher, particularly in relation to the contributions of their academic and industrial co-authors and co-inventors. We find clear linkages, and varied degrees of knowledge transformation, between the technologies in their present form and long-past outputs of the individual, via the publications of the inventor and the literature cited by the patent applications. We also find that the individual demonstrates a high level of absorptive capacity, incorporating and adapting exogenous knowledge into their own knowledge base.

Conference Topic

Technology and Innovation Including Patent Analysis (Topic 5).
Collaboration Studies and Network Analysis (Topic 6)

Introduction

In innovation research, analyses have encompassed various levels of aggregation and address different aspects. For analyses concerning knowledge transfer mechanisms, when examining the minutiae of mechanisms and mediums (such as those of tacit or codified knowledge, R&D networks, formal or informal collaborations), difficulties arise. These difficulties stem from enormous complexities of the knowledge involved in the science and related technologies. The end technological object is the result of the knowledge input and accretion over time into a coherent, and critical, mass. We elaborate upon a method by Gurney *et al* (2012) to discern the knowledge contributions of a specific inventor/author to a patent corpus and the technologies they represent. We utilise two of the output indicators typically used in this and other studies, those of patents and publications. The concepts and practices embodied and codified in the publications and patents were linked to each other, through the citations to literature found in the patent documents. Through linking the two corpora of knowledge the actual knowledge contributions to the development of an idea from inception to product were demonstrated.

The core of this paper discusses the multiple aspects of absorptive capacity, knowledge transfer and transformation, including how scientific knowledge is incorporated into practices, skill sets and eventually artefacts. We then discuss the context and history of our test case. Following this, we briefly summarise the methodology, along with descriptions of the indicators we use followed by the visualisation and clustering techniques employed in our analysis. Our results and conclusions follow, ended with our discussion and implications for further analyses and policy.

Conceptual Framework

The most common and widely cited knowledge transfer mechanisms and inputs are patents, publications, informal and formal interactions, personnel hiring, licensing, R&D collaborations, contract R&D and consulting (Cohen, W.M. et al., 2002). With each of these mechanisms the medium of knowledge transfer can be either codified (such as, for example, patents and publications) or tacit (such as, for example, R&D collaborations and personnel hiring). Key to the reception and implementation of these mediums is the absorptive capacity of the unit under study.

The organisational infrastructure required for facilitating the development and transfer of knowledge depends heavily on the recipient knowledge platform. The knowledge assets (Nonaka, 1994), sector roles (Baba et al., 2009) and older science-push and demand-pull concepts (Langrish et al., 1972), factor into the knowledge base's receptivity. This receptivity is known as 'absorptive capacity' (Cohen, W. M. & Levinthal, 1990) and can best be described as "[t]he ability of a firm to recognize the value of new, external information, assimilate it, and apply it to commercial ends is critical to its innovative capabilities," (p.128).

On an individual level, select individuals act as gatekeepers, such as star (Zucker, L. G. & Darby, 1996) or core (Furukawa & Goto, 2006) scientists. The concept of absorptive capacity has been expanded on significantly by Zahra & George (2002) to include potential and realised absorptive capacity and address (1) Acquisition – the role of prior knowledge or capabilities and the infrastructure already in place; (2) Assimilation – exogenously generated knowledge needs to be understood prior to incorporation; (3) Transformation – the ability to meld exogenous and endogenous knowledge, to create novel fundamental or applied knowledge and (4) Exploitation – the usage of novel knowledge generated during transformation.

Patents have been used as indicators (Schmookler, 1966) for multiple purposes (e.g. Griliches (1998), Schmoch (1993) and Fleming (2001)) as they are highly detailed evidence of technological progress (Tijssen, 2002). Some drawbacks exist, for example, not all innovations are patented (Arundel, 2001; Arundel & Kabla, 1998) or some innovations are kept secret (Brouwer & Kleinknecht, 1999). Publications serve as the primary indicators for the defining characteristics and development of science. They are the most visible outcome of scientific endeavours, and an extensive range of indicators and methodologies have been developed. Analyses using patents or publications are typically based around the meta-data e.g. Title words, abstract words and keywords (Courtial et al., 1993; Engelsman & van Raan, 1994), patent classifications (Leydesdorff, 2008; Tijssen & Van Raan, 1994), and publication/patent citations (Karki, 1997; Meyer, M. S., 2001).

Citation studies using patent-to-literature citations (Meyer, M., 2000; Meyer, M. S., 2001; Meyer, M., 2002; Narin, 1976, 1994) typically rely on direct citation linkages. Non-patent literature references (NPLRs) exhibit different characteristics based on their source, who includes the reference, the patenting offices and completeness of inclusion (Criscuolo & Verspagen, 2008) and their scientific-ness (Callaert et al., 2006). NPLRs from applicants or examiners have typically been treated as being of differing importance (Karki, 1997) but we choose to utilise both types as the presence of citations to literature in patent documents indicates a cognitive link to, or awareness of, the related scientific concepts (Tijssen, 2001), no matter the source of the NPLRs.

University-based scientists publish primarily to extend their professional and intellectual prowess and regular publishing is considered a requirement. There has been an increase in the rate of university patenting linked to institutional and national level changes (Owen-Smith & Powell, 2003; Zucker, L. G. & Darby, 1996), and the increased interest in academic spin-offs and spin-outs (Owen-Smith & Powell, 2003; Zucker, L. G. & Darby, 1996; Zucker, L.G. et al., 1999). With firm-based publishing efforts, the firm stands to gain (or lose) more from the publication process than the author, such as – higher rates of approval of patents (McMillan et al., 2003), a window and source into various fields (Schartinger et

al., 2002) and to stronger ties with future progenitors of knowledge (Hicks, 1995; Zucker, L. G. & Darby, 1996).

Case selection

Our case study involves a prominent Japanese biotechnology researcher, Professor Yusuke Nakamura, who is heavily involved in cancer therapeutics at the University of Tokyo, where he was head of the Human Genome Center. Nakamura founded OncoTherapy Science Inc. (OTS) in April of 2001 to research and develop anti-cancer medicine, cancer therapy and cancer diagnosis based on oncogenes and proteins. He maintains direct links between his research at the University of Tokyo and research conducted at OTS allowing us to draw upon his extensive publishing history as well as his numerous patenting activities, both at the University of Tokyo and OTS.

Method

Data collection

The sources and type of data come from (1) Patents – all patent applications with OncoTherapy listed as an applicant were extracted from the EPO PatSTAT database (2000-2008) with all inventors; (2) Publications – all publications with OncoTherapy listed as an institution were downloaded from WoS (all up to 2011); and all publications with Nakamura listed as any of the authors. These base data were parsed using SAINT (2009) and managed in a relational database. Further data were collected from the patents – specifically (where found) (a) In-text non-patent literature references (IT-NPLRs) and (b) Bibliographic NPLRs (B-NPLRs). The patent documents were grouped by INPADOC family and the associated data aggregated to the parent INPADOC family with each collective representing a specific technology (Martinez, 2010). Where possible the NPLR were identified and matched to their ISI WoS twins and added to the extant set. The origins of each document within the combined set were recorded.

The similarities between publications (both NPLR and Nakamura's) were calculated based on their shared cited reference and title word combinations (van den Besselaar & Heimeriks, 2006). A network was constructed using the publications as nodes and the edges representing the degree of similarity as calculated above. The research streams of publications within the network were assigned by utilising a community detection algorithm developed by Blondel et al (2008). Once the initial research stream assignment was completed, the general streams were isolated and the community detection algorithm was run again to produce smaller concept clusters.

The INPADOC families were clustered using the International Patent Classifications (IPC) codes, the use of which for indicators of knowledge-relatedness has been well-developed (Breschi et al., 2003; Jaffe, 1986).

The NPLRs were co-located within the general research streams based on the level of similarity of shared title word and cited reference combinations. By linking the INPADOC families to the general publication communities in which their NPLRs are co-located, we can infer that there is at least a degree of shared knowledge features between the publication community and the citing INPADOC families.

For more specific knowledge features, the second layer of concept clusters provided a finer-grained view into the communities. Within each concept cluster, the source composition of publications varies. In our case study, in which Nakamura is the primary producer of the publications, each concept can potentially contain a mixture of publications authored by Nakamura and either cited or not, and NPLR not authored by Nakamura. Varying proportions of source publications imply differing levels of imparted or similar knowledge features of the publications. Where Nakamura is not cited but his publications are highly similar, we assume similar skillsets and familiarity of topics and processes of the research. With a concept cluster containing both NPLR and non-NPLR publications by Nakamura, this implies direct contributions of the concepts researched and implemented skill sets. Where there is a combination of all three types, we assume there are direct contributions to concepts and skill sets, and a shared knowledge base and minimum required skill sets.

To visualise the publication community structure over time, we employ a method introduced by Horlings & Gurney (2012) where cognitive communities or research trails over time are transformed based on the time ranges of each community to latitude and longitude coordinates to be displayed on an equirectangular map.

Knowledge capture mechanisms

Following on Zahra & George's (2002) dimensions of absorptive capacity (acquisition, assimilation, transformation and exploitation), we are able to examine in detail: (1) the reputational and applicability aspects of the scientific base work (Hullmann & Meyer, 2003) conducted by Nakamura; (2) the markers for what other fields of science are being utilised by the technologies (Karki, 1997; Schmoch, 1993); (3) the degree of shared knowledge features (such as concepts, knowledge bases and, to a certain extent, skill sets); (4) the level of input from co-inventors of Nakamura; (5) and if Nakamura incorporated skill sets acquired during the development of the technologies and applied them to further his fundamental scientific research by knowledge creation feedback (Fischer, 2001; Tijssen, 1998).

Results

Patents and patent families

In total we collected 242 patent application documents via PatSTAT (Oct 2011) with Nakamura listed as inventor and OncoTherapy as assignee. The patent documents came from 90 INPADOC families, and were composed of 115 priority patents. The earliest patent filing date was March 2000, and the latest was November 2008. The maximum, minimum, average and median numbers of patent applications per INPADOC family are, respectively, 23, 2, 5.3 and 4.

Clustering of INPADOC families by IPC

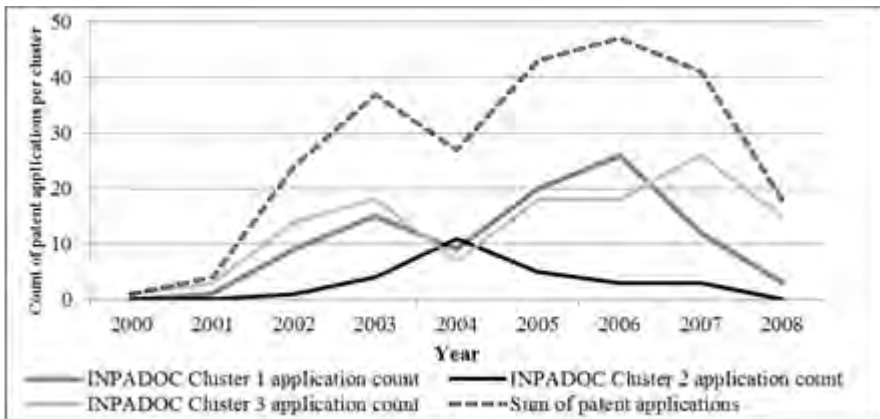


Figure 1(a) INPADOC cluster patent count.

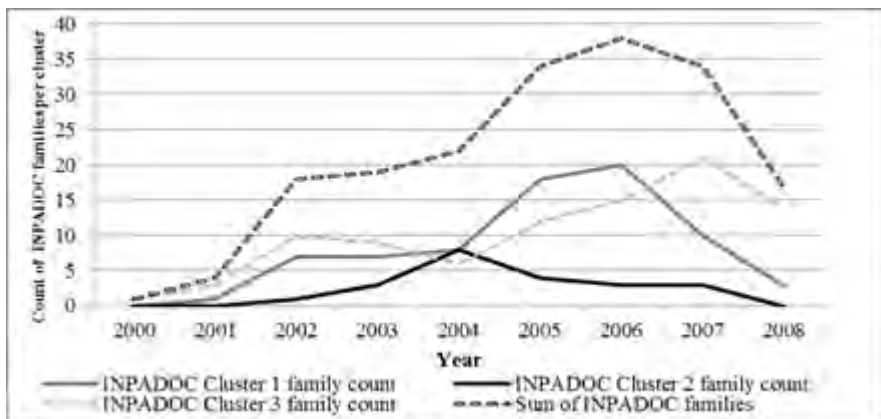


Figure 1(b) INPADOC cluster family count.

Three primary INPADOC clusters were found, using main group IPC data. The growth in the number of patent applications and INPADOC families per cluster are shown in Figures 1(a) and (b). In 2002 and 2004, the number of unique INPADOC

families increased at a slower rate suggesting a period of specialisation within OncoTherapy. From 2004, the increased application rates and increased number of unique families suggest a diversification period. Between 2002 and 2004, Clusters 1 and 3 (dark grey and light grey lines respectively) displayed specialisation whilst Cluster 2 (black border) tended to diversification. In 2004, Cluster 2 peaked and tended to specialisation, whilst 1 and 3 showed overall decreases.

The 2-mode network in Figure 2 demonstrates the specific areas shared by each INPADOC cluster and also serve to highlight which clusters have specialised technological areas that are only applicable to each cluster. As shown in Figure 2, the primary areas at the main group IPC levels addressed by the INPADOC clusters relate primarily to the use of micro-organisms, enzymes, peptides and growth factors, recombinant DNA technologies and medicinal preparations using the peptides and RNA.

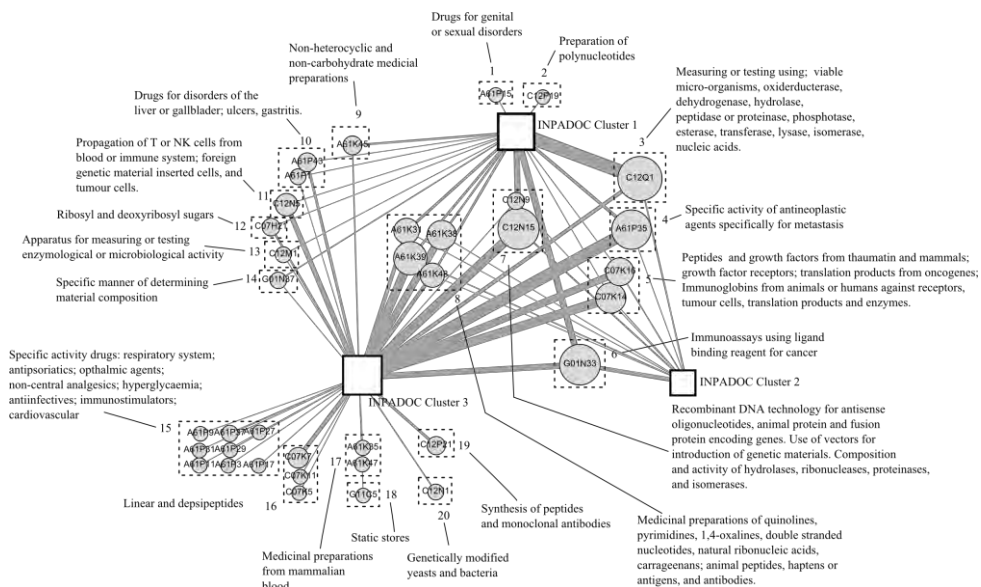


Figure 2 Annotated 2-mode network of main group level IPC and inpadoc family clusters. (Note: Node inpadoc clusters= count of inpadoc families, node size main group IPC nodes=count of patent applications citing main group IPC code. Edge weight=proportional count of number of patent applications utilising the main group IPC code.)

Publications, NPLRs and patents

Nakamura has published a large number of publications with 931 publications over 33 years. His first publication was in 1977 and at a rate under 5 per year until 1987. Between 1988 and 1994, he published between 5 and 10 publications a year and at present, he (co)publishes at a rate of 50 a year.

In total we were able to positively link 525 unique occurrences of B- and IT-NPLRs to the 242 patent applications. Of these NPLRs, 147 were uniquely B-NPLRs, 313 were uniquely IT-NPLRs and 65 NPLR were shared. The most cited NPLR is cited by 41 different patent applications. The most cited publications come from the time period of 1996-2004 with less than 10% of NPLR citations going to publications older than 1996.

Table 1 summarises the distribution of NPLR, the content of each stream, and the links to INPADOC clusters. Figure 3 shows the similarity network of the publications and NPLRs over time. Most Nakamura-authored NPLR are located in streams 7 and 13 and the bulk of patent citations are to Streams 9 and 13.

Table 1 Publication stream summary.

Stream	Total (Nakamura/NP LR /Both)	Start	End	Summary	INPADOC clusters
1	157(84/73/0)	1978	2011	Cell biology, nuc. acids, proteins, polypeptides, factor regul.	1, 2, 3
2	273(182/90/1)	1978	2007	Gene-mapping, novel genes, human genes	1, 2, 3
3	2(0/2/0)	1979	1987	RNA	2, 3
4	85(5/80/0)	1987	2008	Cancer gene expression	1, 2, 3
5	169(133/35/1)	1987	2009	Breast cancer, gene mutation	1, 2, 3
6	2(2/0/0)	1988	1989	Mouse liver	-
7	135(97/18/20)	1988	2011	Gene expr., cancer (prostate, liver, pancreas), therap. targets	1, 2, 3
8	15(0/15/0)	1988	2005	Endocrinology, mouse-human models, porcine spinal-cord	2, 3
9	78(6/72/0)	1989	2007	Lymphocytes, melanomas, peptides, antigens	1, 2, 3
10	8(0/8/0)	1991	2002	Endometriosis, fertility and sterility	3
11	6(5/1/0)	1992	1995	Pharmacology, analogs, glycines	2
12	15(0/15/0)	1993	2005	Methylation (histone and glycine)	2, 3
13	159(110/16/33)	1994	2010	Gene expression, cdna microarrays	1, 2, 3
14	8(6/2/0)	1996	2005	Phospholipase, cell receptors	2
15	15(15/0/0)	1996	2005	OLETF rats, diabetes	-
16	20(20/0/0)	1997	2003	Congenital disorders	-
17	2(0/2/0)	1998	2001	Hepatology	2, 3
18	183(183/0/0)	1999	2011	Japan and population specific cancers	-
19	2(0/2/0)	1999	2000	NFAT mechanisms and inhibition	2

Co-inventors and partner institutes

Figure 4 shows the distribution of Nakamura's co-inventors in the publication corpus. Many publications are authored with Nakamura's co-inventors, with some publications cited as NPLR where Nakamura is not an author. This would seem to

indicate that the knowledge utilised by the patent applications stems not only from Nakamura, but also from his co-inventors. However, the relative scarcity of cited NPLR without Nakamura as author but with one of his co-inventors authoring would suggest that the knowledge comes from within Nakamura's research group.

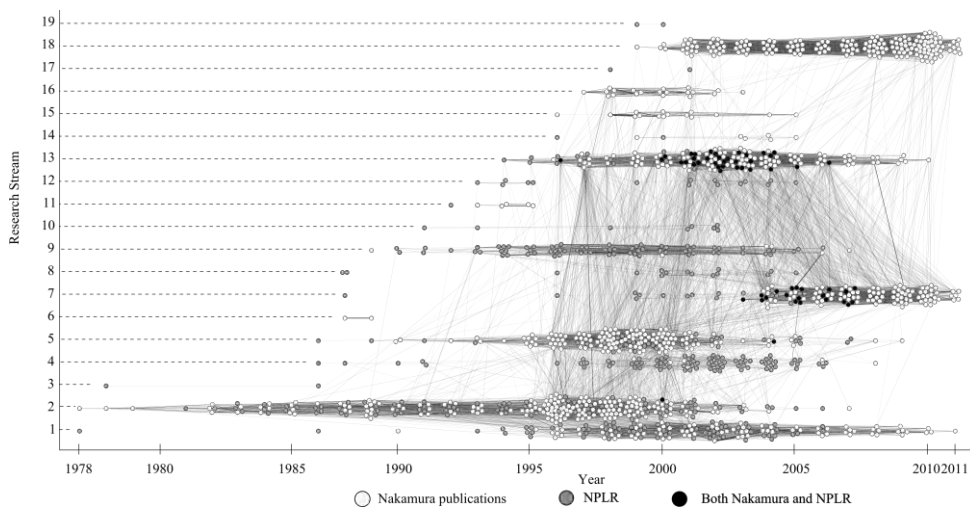


Figure 3 Longitudinal and research stream clustering of Nakamura and NPLR publications. (Note: edges=degree of title word/reference combination similarity. Node colour=source where white=Nakamura publications, Grey=NPLR, Black=Both Nakamura and NPLR)

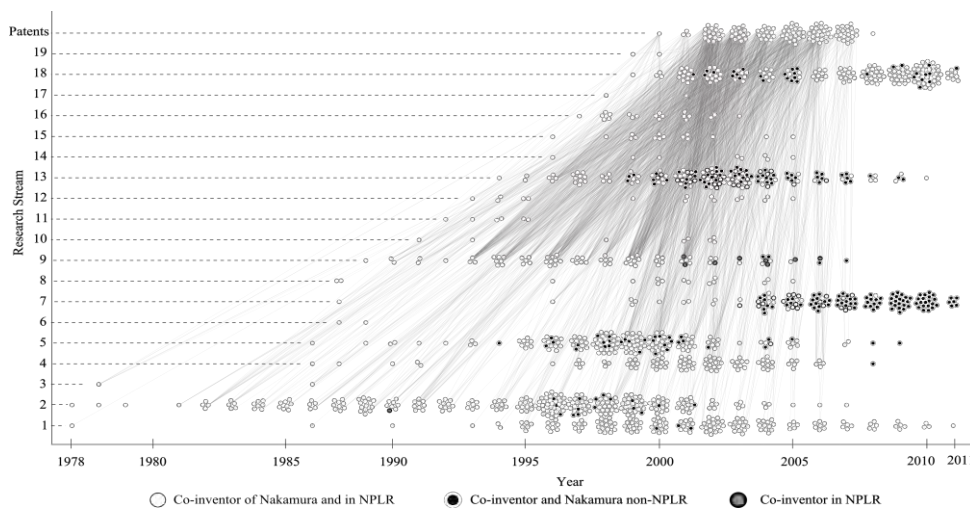


Figure 4 Co-inventor location in research streams. (Note: Only edges between patent applications and cited publications are shown (both IT-NPLR and B-NLPR)).

Within the 77 INPADOC families on which OncoTherapy is listed as assignee and Nakamura as inventor, Nakamura has 10 recurring co-inventors, with 4 of these co-inventors also patenting without Nakamura. OncoTherapy has 6 researchers that patent without Nakamura, but the vast majority of INPADOC families primarily stem from patent applications with Nakamura listed as inventor.

OncoTherapy collaborates on patents with only two organisations, the University of Tokyo in 26 different INPADOC families, and Sentan Kagaku Gijutsu Incubation Center in one INPADOC family. The University of Tokyo is present in just under a third of OncoTherapy's INPADOC families, which, considering Nakamura is based at the university, is not particularly high. The fact that, overall, there is only 1 significant patenting organisational partner for OncoTherapy's technologies is interesting.

Concept clusters

From the 19 research streams, we extracted 66 concept clusters (CCs) that contain NPLRs (both non-Nakamura- and Nakamura-authored). We linked these CCs to the citing INPADOC families and the designated INPADOC clusters. Presented in Figures 5 (a)-(c) are citations to CCs from the INPADOC clusters. Due to space constraints, we have chosen to focus on streams 1, 7, 9 and 13 and their CCs.

From Figure 5(a) – containing only NPLR not authored by Nakamura thus outside Nakamura's expertise, the INPADOC clusters rely heavily, and from an early stage, on CC 9/0 and CC 9/1 (research related to the cytotoxic effect of lymphocytes, and human leukocytes and antigens). INPADOC cluster 1 exclusively cites research from CC 7/2 (increasing rates of bile duct cancer) and CC 1/2 (mRNA binding proteins expression and cancer proteins).

Figure 5(b) shows CCs containing both non-Nakamura-NPLRs and non-NPLR-Nakamura publications. This combination of sources indicates that there is some immediate similarity between research performed by Nakamura and the cited publications. In many cases, the research is cited from an early stage (as seen by the grey edges between nodes) but there is a fair degree of research cited later in the technologies' development phases (dashed and solid black edges). CCs 9/2, 9/3 and 9/4 are cited early by all three clusters, and Nakamura only starts to publish much later in these topics (also seen in Figure 3).

All three INPADOC clusters cite research in CCs 1/4 and 1/5, but again Nakamura's publications related to those topics are only published later. For CC 1/1, cited exclusively by INPADOC cluster 1 in the middle phase of its development, Nakamura - whilst having published extensively in that concept cluster – is not cited at all.

Figure 5(c) shows the CCs considered to contain the most specific aspects of research performed by Nakamura. In most cases, the INPADOC clusters cite the CCs from an early stage but in many cases Nakamura only published later in these topics. This is a strong indicator that Nakamura recognized the necessity of the knowledge in these CCs to further develop the technologies, and assimilated and transformed the content for future research purposes.

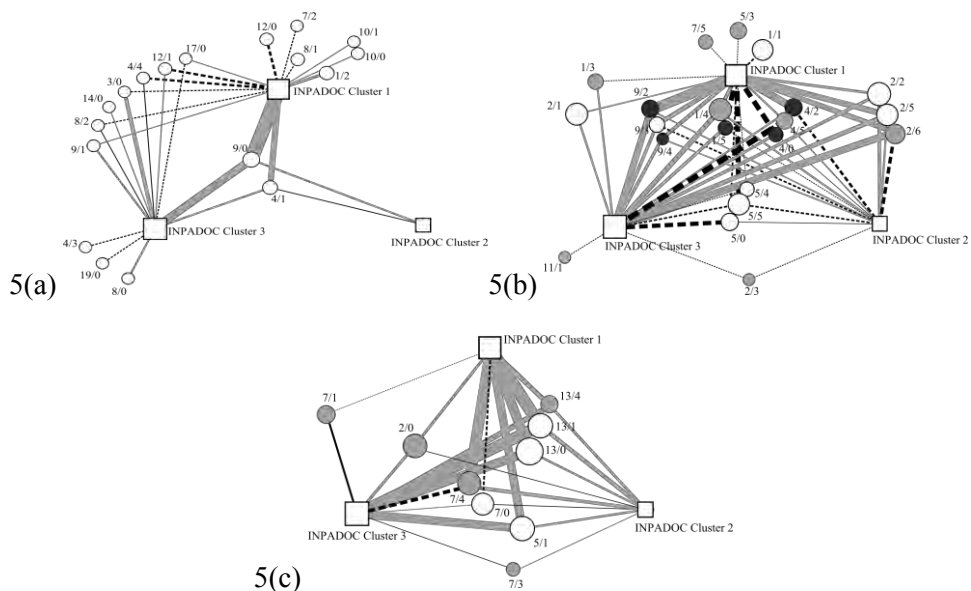


Figure 5 (a)-(c) Concept clusters cited by inpadoc clusters containing (a) only NPLR not authored by Nakamura; (b) NPLR not authored by Nakamura and publications by Nakamura not cited by the patent applications and (c) NPLR authored by Nakamura (Note: For concept labels, a/b, a=parent stream ID, and b=concept ID. Size of nodes=count of publications or count of inpadoc families. Thickness of edges =number of citing inpadoc families. Edge colours: age of the inpadoc cluster the concept is cited, grey=early, dashed=middle, black=late. CC node colours for (b) and (c): White=Nakamura publications present from start, gray= Nakamura publications present from middle time period, black= Nakamura publications present at end of time period))

Summarising, in stream 1, Nakamura publishes extensively but is not cited by the patent applications at all. The degree of exogenously-generated knowledge is high, with no direct contributions by Nakamura. However, the shared knowledge base and shared minimum skill set is significant as only one of the five CCs cited do not contain any Nakamura publications.

With stream 7, initially the INPADOC clusters barely cite the stream at all. Up to 2004 the first cited NPLRs were all non-Nakamura NPLRs but from 2004 onwards Nakamura publishes prolifically and is often cited. The proportionally large number of Nakamura-NPLRs and Nakamura's knowledge base and skill sets are now integral to the technologies.

The technologies cite stream 9 extensively but Nakamura's role is limited. He is not directly cited but does publish at later stages in all of the cited CCs. In short, the necessary scientific aspects derived from stream 9 are exogenously sourced. However some of the topics relate to background information.

Nakamura-authored publications dominate stream 13 with a third of his publications cited by the technologies. In one CC (13/4) Nakamura is not the first

to publish, with some NPLRs coming from others. The role of Nakamura's research in stream 13 and its contributions to the technologies of clusters 1-3 is more obvious as the publications in this stream are authored almost entirely by Nakamura.

Summary and conclusion

Considering the enormous volume of data available with our approach, we chose to focus on four specific streams of publications and their impact (through citation links and topic similarity) on the patent applications. We also reduced the specificity of the technologies by aggregating the patent applications into INPADOC families and then further into INPADOC clusters. At an obvious loss of detail, we feel that the aggregation was necessary to better analyse the knowledge and skillset contributions of Nakamura as an individual.

Nakamura's impact within these four streams on the INPADOC clusters was viewed through the lens of the adoption and adaptation aspects of Zahra and George (2002) and their respective source of knowledge, be they exogenously or endogenously generated.

Acquisition – this dimension primarily details the role of prior knowledge or capabilities and the infrastructure already in place. The first step in this aspect is recognising knowledge that is or would be useful to the development of the technologies. By examining the degree of required knowledge through Figures 4 and 5(a) we gain insight to this aspect. Co-inventors are considered here as they provide necessary expertise and skillsets.

Assimilation – By conducting research in the topic areas required for the technologies, whether through a non-concerted approach or a cumulative directed approach, the codified and tacit skills and insights developed directly impact the development of the technologies. In this sense, the process of 'learning-by-doing' seems to be prevalent. In examining Figures 5(a) and (b) we can see the specific topics and levels of contribution by Nakamura and at what stages of the science his contributions become visible.

Transformation – addresses the ability to meld exogenous and endogenous knowledge, to create novel fundamental or applied knowledge. Streams 7 and 13 from Figure 3 provide examples of this. Taking into consideration the degree of similarity between stream 5 and stream 13, we see a strong link, particularly around 1996 and 2000 coinciding with bursts of publishing one year later in stream 13. The skill sets and knowledge acquired in practising research in topics within stream 5 have had a significant influence on the required knowledge and skills sets for stream 13. The same behaviour can be discerned between streams 13 and 7, where stream 13 provides the required knowledge and skill sets for the topics in stream 7. Translating this to the patent applications: where the technologies previously relied on exogenously generated knowledge from streams 7 and 13, the endogenously generated knowledge of stream 5 was successfully acquired, assimilated and transformed for use in streams 13 and 7.

On a methodological level, our approach benefits from its ability to encompass both the macro and micro views. Our approach can isolate and highlight specific aspects of utilised knowledge in relation to the knowledge features already in place. We are able to co-locate the knowledge features of individuals who contribute to the publications and patent applications, not through the direct citations of NPLRs, but through the co-location of NPLRs in the environment.

A disadvantage of our method as outlined above is the complexity of the process. Due to this complexity we chose to aggregate the technologies into clusters of INPADOC families. This limits our attention to detail within the technologies but allows a thorough examination of the contributions of an individual (in our case Nakamura). The possibility exists to aggregate on the publication side and examine in detail the characteristics of the technologies being produced.

We see this method aiding in the evaluation of technologies and the contributions of those involved in the development of the technologies. With the addition of funding information in the meta-data extracted from WoS it would then be possible to trace the results of such funding to its exploitation phase and results. The scaling up of this method would allow research groups, departments or entire research institutes or infrastructures to map their contributions in the early stages of the development of a technology right through to their exploitation or implementation. This would be invaluable to funding agencies and universities for reporting on their research achievements, as in many cases the end-point of fundamental and applied research may be so far removed from the origin as to be unrecognisable.

Acknowledgements

We gratefully acknowledge the support of a JSPS/CNRS Grant (DP) and Aix Marseille University (KS). The authors would like to thank the two anonymous reviewers for their useful comments and input.

References

- Arundel. (2001). The relative effectiveness of patents and secrecy for appropriation. *Research policy*, 30(4), 611-624.
- Arundel, & Kabla. (1998). What percentage of innovations are patented? Empirical estimates for European firms. *Research policy*, 27(2), 127-141.
- Baba, Shichijo, & Sedita. (2009). How do collaborations with universities affect firms' innovative performance? The role of. *Research Policy*, 38(5), 756-764.
- Blondel, Guillaume, Lambiotte, & Lefebvre. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008, P10008.
- Breschi, Lissoni, & Malerba. (2003). Knowledge-relatedness in firm technological diversification. *Research policy*, 32(1), 69-87.
- Brouwer, & Kleinknecht. (1999). Innovative output, and a firm's propensity to patent.: An exploration of CIS micro data. *Research policy*, 28(6), 615-624.

- Callaert, Van Looy, Verbeek, et al. (2006). Traces of prior art: An analysis of non-patent references found in patent documents. *Scientometrics*, 69(1), 3-20.
- Cohen, & Levinthal. (1990). Absorptive Capacity: A New Perspective on Learning and Innovation. *Administrative Science Quarterly*, 35(1, Special Issue: Technology, Organizations, and Innovation), 128-152.
- Cohen, Nelson, & Walsh. (2002). Links and impacts: the influence of public research on industrial R&D. *Management Science*, 1-23.
- Courtial, Callon, & Sigogneau. (1993). The use of patent titles for identifying the topics of invention and forecasting trends (Vol. 26, pp. 231-242): Springer.
- Criscuolo, & Verspagen. (2008). Does it matter where patent citations come from? Inventor vs. examiner citations in European patents. *Research policy*, 37(10), 1892-1908.
- Engelsman, & van Raan. (1994). A patent-based cartography of technology. *Research policy*, 23(1), 1-26.
- Fischer. (2001). Innovation, knowledge creation and systems of innovation. *The Annals of Regional Science*, 35(2), 199-216.
- Fleming, & Sorenson. (2001). Technology as a complex adaptive system: evidence from patent data. *Research Policy*, 30(7), 1019-1039.
- Furukawa, & Goto. (2006). Core scientists and innovation in Japanese electronics companies. *Scientometrics*, 68(2), 227-240.
- Griliches. (1998). Patent statistics as economic indicators: a survey: University of Chicago Press.
- Gurney, Schoen, Horlings, et al. (2012). *Knowledge Capture Mechanisms in Bioventure Corporations*. Paper presented at the Proceedings of 17th International Conference on Science and Technology Indicators.
- Hicks. (1995). Published papers, tacit competencies and corporate management of the public/private character of knowledge. *Industrial and Corporate Change*, 4(2), 401.
- Horlings, & Gurney. (2012). Search strategies along the academic lifecycle. *Scientometrics*, 1-24.
- Hullmann, & Meyer. (2003). Publications and patents in nanotechnology. *Scientometrics*, 58(3), 507-527.
- Jaffe. (1986). Technological Opportunity and Spillovers of R & D: Evidence from Firms' Patents, Profits, and Market Value (Vol. 76, pp. 984-1001): JSTOR.
- Karki. (1997). Patent citation analysis: A policy analysis tool. *World Patent Information*, 19(4), 269-272.
- Langrish, Gibbons, Evans, & Jevons. (1972). *Wealth from knowledge: a study of innovation in industry*: Halstead Press Division, Wiley.
- Leydesdorff. (2008). Patent classifications as indicators of intellectual organization. *Journal of the American Society for Information Science and Technology*, 59(10), 1582-1597.
- Martinez. (2010). *Insight into different types of patent families*: OECD.
- McMillan, Mauri, & Robert III. (2003). The impact of publishing and patenting activities on new product development and firm performance: the case of the

- US pharmaceutical industry. *International Journal of Innovation Management*, 7(2).
- Meyer. (2000). Does science push technology? Patents citing scientific literature. *Research Policy*, 29(3), 409-434.
- Meyer. (2001). Patent citation analysis in a novel field of technology: An exploration of nano-science and nano-technology (Vol. 51, pp. 163-183): Springer.
- Meyer. (2002). Tracing knowledge flows in innovation systems. *Scientometrics*, 54(2), 193-212.
- Narin. (1976). Evaluative bibliometrics: The use of publication and citation analysis in the evaluation of scientific activity. *Cherry Hill, NJ: Computer Horizons*, 338, 27.
- Narin. (1994). Patent bibliometrics. *Scientometrics*, 30(1), 147-155.
- Nonaka. (1994). A dynamic theory of organizational knowledge creation. *Organization science*, 5(1), 14-37.
- Owen-Smith, & Powell. (2003). The expanding role of university patenting in the life sciences: assessing the importance of experience and connectivity. *Research policy*, 32(9), 1695-1711.
- Schartinger, Rammer, Fischer, & Fröhlich. (2002). Knowledge interactions between universities and industry in Austria: sectoral patterns and determinants. *Research Policy*, 31(3), 303-328.
- Schmoch. (1993). Tracing the knowledge transfer from science to technology as reflected in patent indicators. *Scientometrics*, 26(1), 193-211.
- Schmookler. (1966). *Invention and economic growth*: Harvard University Press Cambridge, MA.
- Somers, Gurney, Horlings, & Van den Besselaar. (2009). *Science Assessment Integrated Network Toolkit (SAINT): A scientometric toolbox for analyzing knowledge dynamics*. The Hague: Rathenau Insitute.
- Tijssen. (1998). Quantitative assessment of large heterogeneous R&D networks: the case of process engineering in the Netherlands1. *Research policy*, 26(7-8), 791-809.
- Tijssen. (2001). Global and domestic utilization of industrial relevant science: patent citation analysis of science-technology interactions and knowledge flows. *Research policy*, 30(1), 35-54.
- Tijssen. (2002). Science dependence of technologies: evidence from inventions and their inventors. *Research policy*, 31(4), 509-526.
- Tijssen, & Van Raan. (1994). Mapping changes in science and technology. *Evaluation Review*, 18(1), 98-115.
- van den Besselaar, & Heimeriks. (2006). Mapping research topics using word-reference co-occurrences: a method and an exploratory case study. *Scientometrics*, 68(3).
- Zahra, & George. (2002). Absorptive capacity: A review, reconceptualization, and extension. *Academy of management review*, 185-203.

- Zucker, & Darby. (1996). Star scientists and institutional transformation: Patterns of invention and innovation in the formation of the biotechnology industry. *Proceedings of the National Academy of Sciences of the United States of America*, 93(23), 12709.
- Zucker, Darby, & Brewer. (1999). *Intellectual capital and the birth of US biotechnology enterprises*: National Bureau of Economic Research.