



Research article

Evaluating “startup readiness” for researchers: case studies of research-based startups with biopharmaceutical research topics

Tomotaka Goji^{a,*}, Yuki Hayashi^a, Ichiro Sakata^b^a The University of Tokyo Edge Capital Partners, Co., Ltd. (UTECE), Tokyo, Japan^b Innovation Policy Research Center, Institute of Engineering Innovation, School of Engineering, The University of Tokyo, Tokyo, Japan

ARTICLE INFO

Keywords:

Business
 Computer science
 Pharmaceutical science
 Biotechnology
 Startup readiness
 Research-based startups
 Paper citation networks
 Co-authorship networks
 Academic entrepreneurship
 Startup finances
 Venture capital
 User-inspired fundamental research
 Technological innovation
 Biopharmaceuticals

ABSTRACT

For research domains such as life sciences, which pursue fundamental scientific understanding and applications intended for immediate use, academic entrepreneurship has played a pivotal role in commercialization. This paper presents an evaluation method of researchers related to user-inspired fundamental research, using global databases of startup finances and academic research papers of “startup readiness.” Case studies of startups related to biopharmaceutical research topics suggest that the biopharmaceutical field has rich opportunities stemming from scientific research, commercialization, and entrepreneurship. This evaluation method sorts specific industry segments by which financing activities are active, and by which related growing research topics attract increased academic attention. We constructed networks of author citation and co-authorship from paper citation networks related to research topics in industry segments in the biopharmaceutical domain. Results obtained across all research topics we surveyed demonstrated that authors in the top 10% of degree centrality ranking in both networks are far more likely to be startup participants than other authors. Our computational approach might provide convenient, dynamic, global, and real-time understanding of the “startup readiness” of researchers working with research topics for which academic attention is emerging in actively financed biopharmaceutical fields.

1. Introduction

Conventionally, research examining factors contributing to creation of scientific research-based startups and academic entrepreneurship has specifically examined factors other than scientific research itself. One example of a generally accepted notion about commercialization from advanced technology research is the technology readiness level (TRL) concept. In the mid-1970s, NASA introduced TRL: a criterion to evaluate the maturity of technologies derived from science based on a scale of 1–9, with 1 being the most basic technology and 9 being the most mature technology. The criterion has been used to explain why some new technologies engender industrial transitions. However, it has been used for project management at stages of performance assessment to schedule to budget without addressing the emergence of the research itself [1]. Actually, TRL has been expected to facilitate transdisciplinary expertise between academia and practitioners by supporting the analysis and design of an industry's transition [2].

Patents, which represent practical applications of technology concepts that are classified as TRL 2 according to the TRL concept, have been

regarded as an important antecedent for academic entrepreneurship in life sciences [1, 3]. University scientists who appear as inventors on patents presumably have more opportunities to enter the private sector than other scientists. Such might be the case also for biotechnology startups. By contrast, a research paper itself has not been regarded as a precursor for academic entrepreneurship because the purpose of a research paper has been regarded as communicating research findings to the relevant scientific community and to the general public. Moreover, patents are legal documents used to prevent others from commercializing what research papers describe, thereby facilitating commercialization by patent holders of their intended applications.

The “startup readiness” concept developed in our earlier study [4] was proposed as a criterion to be applied earlier than TRL, focusing on individual scientists. Whereas TRL is used to assess the maturity of a technology, startup readiness is the state at which a researcher is ready to participate in a startup. Given the growing interest in academic startups among increasing stakeholders, academic startups have recently been gaining easier access to venture capital and managerial talent. Long before TRL matures, research topics and researchers show startup

* Corresponding author.

E-mail address: tomotaka.goji@gmail.com (T. Goji).

readiness, which might later beget important firms by leveraging their scientific strengths. These research topics and researchers are investment opportunities for venture capitalists. They represent career opportunities for managerial talent as well. An earlier report of the literature has described that biotechnology research paper publications linked to patents are cited more than publications without a patent link [5], suggesting that research papers citations signal commercial value. We hypothesized that “startup readiness” of research topics and researchers, in contrast to “technology readiness,” can be signaled by research papers. This study presumes the resource-based view of firms (Barney (1991); Kogut and Zander (1992); Conner and Prahalad (1996); Grant (1996)) and its extended literature related to academic startups (Landry et al. (2006); Knockaert et al. (2010); Rasmussen and Borch, (2010); Huynh et al. (2017); Corsi et al. (2019)) to assume that, as with entrepreneurs, startup readiness by academic researchers will increase when either the resources or their coordination will be appropriate or sufficient [6, 7, 8, 9, 10, 11, 12, 13, 14]. Specifically for factors contributing to new academic firm creation, researchers have assessed individual and non-individual determinants of academic entrepreneurship using surveys of scientists classified according to their research protocols. For instance, Rothaermel et al. (2007) report that university policy, faculty, technology transfer offices, investors, founding teams, networks in which a firm is embedded, and other external conditions affect new firm creation [15]. Bercovitz and Feldman (2008) examine faculty members' backgrounds and work environments and their subsequent academic entrepreneurship activities. They find that participation of researchers, if they had accepted the new initiative and had been active in technology transfer, is more likely at institutions where they trained [16]. Jain et al. (2009) investigate the sense-making process accompanying university scientist participation in academic entrepreneurship and potential modification of their role identity. They suggest that scientists participate to preserve their academic role identity [17]. Clarysse et al. (2011) examine how academic professionals' opportunity recognition capacity and their prior entrepreneurial experience affect the likelihood of their involvement in starting up a new venture and shape the social environment and roles of university technology transfer offices [18]. Abreu and Grinevich (2013) analyze determinants of academic engagement such as demographic factors (seniority and gender), research type, entrepreneurial experience and training, and institutional support [19]. Aldridge et al. (2014) examine motivations for scientists to start companies, specifically examining roles of scientist characteristics including academic rank, experience, networks and industry ties, access to human and financial resources, and supportive university conditions [20].

The earlier literature of resource-based theory describes the commercialization of academic research: resources that enable startup creation include knowledge assets, intellectual property assets, financial assets, social capital assets, personal assets, and organizational assets as discussed above. Except for organizational assets that are environmental factors, these assets are individual factors. However, for the biopharmaceutical domain, which is a particularly user-inspired, intense science-based technology commercialization field, we assume that the scientific prominence of researchers is more important than their other attributes. That point has not been explored in earlier studies of this field. This report describes our study assessing the notion that researchers' standing in the biopharmaceutical academic community, which signals their scientific prominence, is a factor indicating their “startup readiness,” while supporting the basic view of resource-based theory.

Furthermore, earlier studies examining factors leading to the creation of startups commercializing scientists' research make little or no reference to any of the following: (1) collection of real-time data related to intended entrepreneurial researchers' academic and startup activities in broad disciplines in a scalable manner (although earlier studies survey past data of scientists in specified academic organizations or regions); (2) selection of specific research topics that entrepreneurial researchers are pursuing (although Abreu and Grinevich (2013) introduce life sciences as research fields with greater commercialization activity) [19]; and (3)

bibliometric analyses of those researchers specifically in terms of their research topics in their research communities (although Rothaermel et al. (2007) and Aldridge et al. (2014) assess academic titles such as professor) [15, 20].

The industry segments presented herein highlight case studies in life sciences, specifically the biopharmaceutical domain. The classic definition of “biopharmaceutical,” both in science and industry, is pharmaceuticals (medicinal products, therapeutics, prophylactics, and in vivo diagnostics) with active agents that are inherently biological in nature and which are manufactured using biotechnology (products manufactured by or from living organisms, usually involving bioprocessing). In addition, “drug” is defined as a pharmaceutical that is inherently chemical (not biological) in nature and which is manufactured using chemical methods. Biopharmaceuticals are distinct from drugs, most of which are composed of small molecules or other synthetic chemical substances. The inherent differences between these two classes include product and active agent sources, identity, structure, composition, manufacturing methods and equipment, intellectual property, formulation, handling, dosing, regulation, and marketing [21]. The biopharmaceutical domain is selected for our case studies here for several reasons. (a) Much of life sciences such as biopharmaceutical research can be characterized as situated in Pasteur's quadrant, a classification of scientific research projects that is aimed at fundamental understanding of scientific problems and at providing immediate benefits for society. Many studies associated with that quadrant have revealed evidence of greater commercialization activities through academic entrepreneurship [22, 23, 24, 25, 26]. In addition, (b) biopharmaceutical science has expanded considerably in terms of both commercialization and entrepreneurship since the beginning of the 21st century. Of the top 10 selling pharmaceutical products worldwide in 2017, 9 were biopharmaceutical; also, 6 had origins in startup companies (Table 1). By contrast, 2001 had only one biopharmaceutical drug without a startup origin [27].

This study explores the proposition that, in user-inspired fundamental research such as that in the biopharmaceutical domain, specific knowledge regarding academic entrepreneurship can be attained as follows: (1) Global databases of startup finances and academic research papers can yield real-time data that are useful to analyze and predict academic entrepreneurship in broad disciplines on a global scale. (2) Data of startup finances and keyword analyses both of the data of startup finances and academic research papers can elucidate trends indicating which scientific topics are becoming active areas of research that enhance “startup-readiness.” (3) Bibliometric analyses of research papers related to the scientific topics above can reveal potentially entrepreneurial scientists with high startup readiness in industry segments related to the scientific topics.

Results of this study contribute to a rich, convenient, and dynamic understanding of academic entrepreneurship. The computational approach used herein can be useful to assess the startup readiness of researchers in emerging research topics because names of research topics and authoring researchers are identifiable easily in real time. Better scalability and adaptability are obtainable using these names instead of published paper titles, personal interviews or field projects. Consequently, our method can be useful even for practitioners including venture capitalists and managerial entrepreneurs who are unfamiliar with science fields, but who aspire to engage in research-based startups. Our approach is also unique because it directly bridges authors in citation networks to a startup financing database, which enables us to relate citation network analysis to the startup and investment contexts. described above.

Concretely, our bibliometric analyses first require construction of author citation networks derived from biopharmaceutical research paper citation networks. Then we refer the author data to the startup database to analyze whether the authors are biopharmaceutical startup participants such as a founder, chief scientific officer, or director. Herein, we present this method of evaluating researchers' startup readiness of biopharmaceutical startups. Then we measure the degree centrality of such

Table 1. Top 10 pharmaceutical products by Global sales in 2017 (million U.S. dollars).

Rank	Product	Therapeutic Subcategory	Vendor Company	Originator	2017 Sales (\$m)
1	Humira	Other anti-rheumatics	AbbVie, Eisai	Knoll	18,923
2	Enbrel	Other anti-rheumatics	Amgen, Pfizer, Takeda	ImmuneX, acquired by Amgen	8,234
3	Revlimid	Other cytostatics	Celgene, BeiGene	Celgene	8,211
4	Rituxan	Anti-neoplastic MAbs	Roche	IDEC Pharmaceuticals, merged with Biogen	7,528
5	Remicade	Other anti-rheumatics	Johnson & Johnson, Merck & Co, Mitsubishi Tanabe Pharma	Centocor, renamed Janssen Pharmaceutical	7,172
6	Herceptin	Anti-neoplastic MAbs	Roche	Genentech	7,126
7	Avastin	Anti-neoplastic MAbs	Roche	Genentech	6,795
8	Eylea	Eye/Ophthalmic preparations	Regeneron Pharmaceuticals, Bayer, Santen Pharmaceutical	Regeneron Pharmaceuticals	6,291
9	Opdivo	Anti-neoplastic MAbs	Bristol-Myers Squibb, Ono Pharmaceutical	Ono Pharmaceutical	5,761
10	Prevnar 13	Vaccines	Pfizer	Wyeth	5,693

Source: Evaluate Ltd. "Top 100 Products in 2024"

authors' nodes in author citation networks and co-authorship networks. Centrality represents the degree to which an author is central in terms of the position in the author citation network. Degree centrality can be ascertained using several methods: often, network analyses are conducted among actors (i.e. network nodes), as some earlier studies have done [28, 29, 30, 31, 32]. Our similar approach measured a centrality index with respect to targeted nodes and others to explain causes of their performance or activities. For an earlier study that analyzed author citation networks, Ding applied PageRank and weighted PageRank algorithms in author citation networks to assess the popularity and prestige of scholars [33]. Our approach differs: Ding analyzed only first authors, whereas we analyzed all co-authors comprehensively in their research field. An earlier study by Cainelli et al. (2015) analyzed co-authorship of economists and co-authorship effects using social network analysis and economic analysis to assess co-authorship structures and to explain researcher productivity in terms of variables including cooperation and two centrality indexes: betweenness and closeness [34].

2. Methodology

This section presents the methodology proposed for this research. The analytical scheme is depicted in Figure 1.

Using the VentureSource global database of startups, we analyzed all financing deals struck between January 1, 2017 and December 31, 2017 to construct a list of industry segments that were most active in venture capital finance in 2017 (A-1). We narrowed this analysis to 2017 while analyzing the years 2014–2017 to identify the most rapidly growing keyword use in research papers and to create author citation networks and co-authorship networks associated with those keywords (A-3 and A-4) because we were interested in how "hot" each industry segment in VentureSource in the latest year.

VentureSource, compiled by Dow Jones & Company, is a comprehensive global database of companies backed by venture capital and private equity in every region, industry, and stage of development. The database includes daily data of startup investment deals according to a specific industry code/subcode. Information related to financing amounts, financing rounds, company overview keywords, startup participants, etc. is available. Using VentureSource, based on the sum of the ranks of both the average financing size and the number of financing rounds (A-1), we compiled a ranked list of the top 30 most active financing industry codes/subcodes among all 281 VentureSource industry codes/subcodes in 2017. Then we extracted industry codes/subcodes for the VentureSource "Biopharmaceutical" industry segment.

For startups in the target codes/subcodes belonging to the biopharmaceutical segment on VentureSource, we surveyed keywords that appeared multiple times (A-2). We analyzed the keyword frequency in research papers of 2014–2017 by searching the Web of Science Core Collection database compiled by Clarivate Analytics. We identified the

keywords showing the most growth in use during that period for each target code/subcode, as emerging research topics (A-3).

Then, we created author citation networks and co-authorship networks based on the research papers for each selected keyword (A-4). Simultaneously, we detected authors in both networks who were participating in startups related to those keywords (A-5).

Finally, after analyzing the distribution of the startup participant authors in the author citation networks and in the co-authorship networks to assess their degree centrality (B-1), we tested the top 10% degree centrality authors of both networks to infer their "startup readiness" (B-2).

A-1. Constructing a list of industry fields most actively financed based on average financing size and number of financing rounds in 2017

Using the VentureSource database, we constructed lists of the most active financing industry fields based on the average financing size and the number of financing deals per field during the 365 days of 2017. We analyzed 2017 to ascertain how financially active each industry segment was during the latest year among the years 2014–2017, throughout which we conducted analyses for A-3 and A-4, as described later. According to VentureSource, 17,681 financing deals were concluded in 2017. We sorted all of their 281 industry codes/sub-codes in descending order by the average financing deal size and by the number of financing rounds. We then constructed our top 30 (approximately top 10%) ranking of most actively financing industry fields, arranged based on the sum of the orders of both the average financing size and the number of financing rounds. We examined the number of financing rounds in addition to the average financing size because this analysis primarily addresses venture capital financing where the number of rounds of financing in seed or early stage companies is important. Private equity investment deals typically have a few large investment rounds in middle or later stage companies (Table 2).

From the top 30 industry fields above, we extracted industry codes/subcodes of the "Biopharmaceuticals" industry segment as case studies for exploration. Particularly, we selected five industry codes/subcodes for an additional survey presented below: "Pharmaceuticals," "Biotechnology Therapeutics," "Immunotherapy/Vaccines," "Small Molecule Therapeutics" and "Gene Therapy."

A-2. Extraction of keywords related to active financing of biopharmaceutical industry fields

Keywords representing company overviews are available in the VentureSource database. From VentureSource, we extracted keywords representing company overviews of the startups with the five industry codes/subcodes described in A-1, by gleaning those keywords from company overviews using the "webdriver" module of Python's

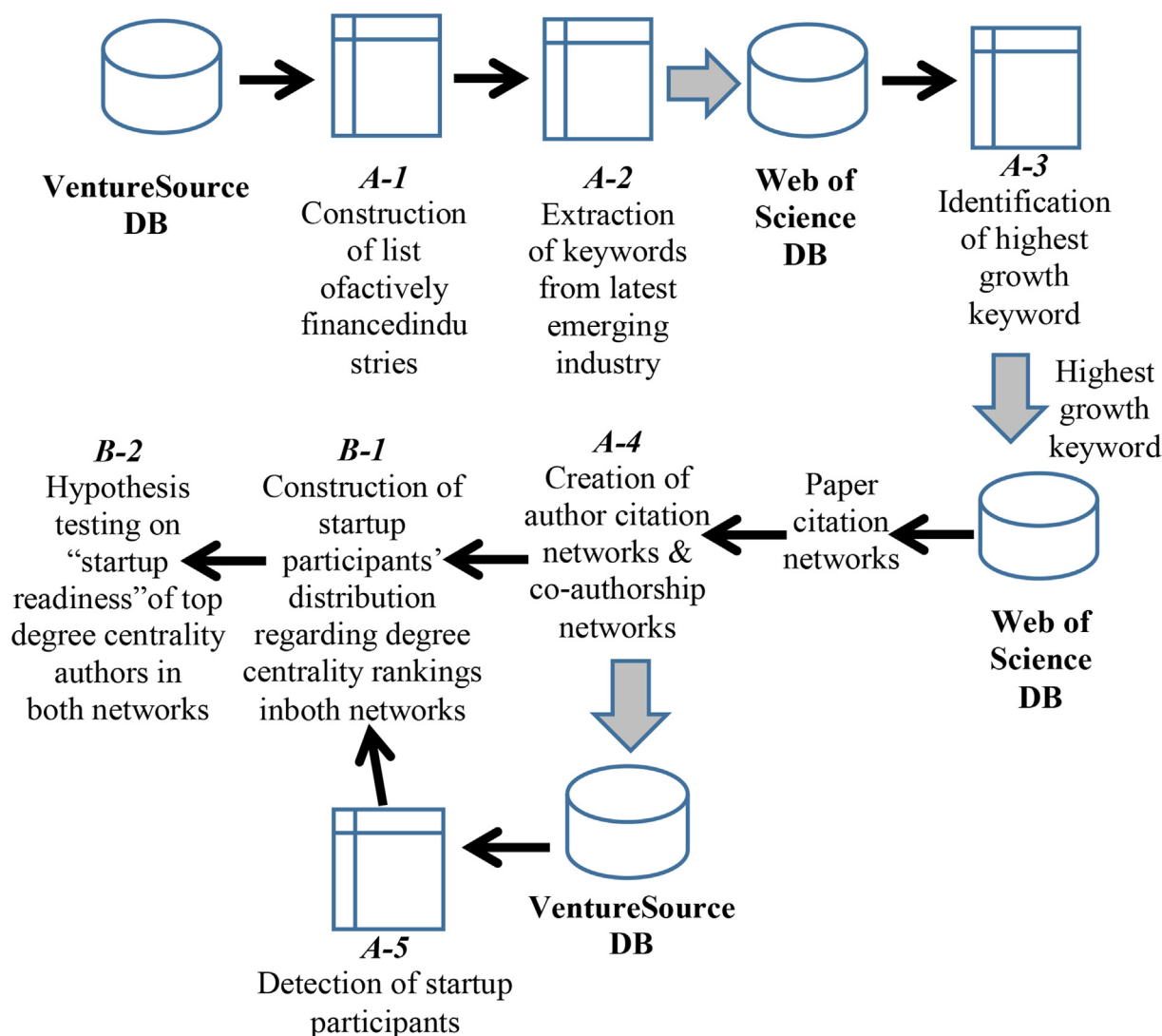


Figure 1. Methodology proposed herein.

“selenium” library. Subsequently, we constructed keyword lists for each industry code/subcode. Keywords appearing twice or more on each list are presented in Table 3, with calculations of each keyword's appearance frequency using Python's “pandas” library.

A-3. Identifying research paper keywords with most usage growth for emerging research topics of actively financing biopharmaceutical industry fields of 2014–2017

Then, we conducted queries of all keywords in Table 3 into the Web of Science Core Collection database to analyze growth in the frequency of those keywords that appeared in the title, keywords, or Keyword Plus of the research papers during 2014–2017. These consecutive years were observed to elucidate the growth of specific research topics throughout what we deemed as an appropriately certain length of time, four years, as determined by reference to leading cases in CRISPR and Cas9, with the maximum years to founding of a startup because its initial associated research paper took four years [4]. Rankings of keywords showing the most growth in incidence for each industry code/subcode were constructed based on growth multiples (Table 4).

From the rankings presented above, we extracted keywords with incidences that more than doubled between 2014 (Start Year) and 2017 (End Year) and which appeared more than 100 times in 2017 (End Year) for an additional survey. Five keywords met these criteria from each

industry code/subcode. All are emerging research topics that have attracted academic attention to a rapidly increasing degree.

2.3.1. Pharmaceuticals: exosome

Exosomes are small microvesicles released from late endosomal compartments of cultured cells in many and eukaryotic fluids, including blood and urine [35, 36]. They are either released from the cell when multivesicular bodies fuse with the plasma membrane, or are released directly from the plasma membrane [https://en.wikipedia.org/wiki/Exosome_\(vesicle\)](https://en.wikipedia.org/wiki/Exosome_(vesicle)) [37]. Moreover, they have specialized functions such as coagulation, intercellular signaling, and waste management [35]. Consequently, growing interest has arisen in their clinical applications. Exosomes might be used for therapy and prognosis, or as biomarkers for health and disease.

2.3.2. Biotechnology Therapeutics: Microbiome

Microbiome refers to ecological communities of commensal, symbiotic and pathogenic microorganisms <https://en.wikipedia.org/wiki/Microbiota> [38, 39] living in and on all multicellular organisms from plants to animals. A microbiome denotes either the collective genomes of microorganisms occupying an environmental niche or the microorganisms themselves <https://en.wikipedia.org/wiki/Microbiota> [40, 41, 42]. By influencing adaptive and innate immune functions, a microbiome can promote or disrupt human health [43].

Table 2. Top 30 most actively financing industry fields among 281 VentureSource industry codes/subcodes based on 17,681 financing deals during January 1, 2017 through December 31, 2017.

Rank	Industry Segment	Industry Code/Subcode	Average Finance Size	Order	# of Rounds	Order
1	Travel and Leisure	Transportation Services	148.25	4	224	18
2	Financial Institutions & Services	Lending	46.31	25	359	5
3	Consumer Information Services	Shopping Facilitators	34.23	36	921	1
4	Business Support Services	Facilities/Operations Management	32.35	43	272	13
5	Consumer Information Services	Email/Messaging	58.61	16	97	52
6	Financial Institutions & Services	Insurance	33.50	38	133	36
7	Financial Institutions & Services	Retail Investment Services/Brokerages	63.95	15	81	60
8	Wholesale Trade and Shipping	Logistics/Delivery Services	29.72	51	166	29
9	Biopharmaceuticals	Pharmaceuticals	32.28	44	129	38
10	Financial Institutions & Services	Payment/Transactional Processing	20.60	86	313	10
11	Business Support Services	Data Management Services	18.50	94	339	6
12	Biopharmaceuticals	Biotechnology Therapeutics	23.87	72	154	32
13	Financial Institutions & Services	Real Estate	23.46	73	140	35
14	Electronics & Computer Hardware	Consumer Electronics	17.71	99	316	9
15	Biopharmaceuticals	Immunotherapy/Vaccines	29.05	54	93	55
16	Medical Devices & Equipment	Medical Lab Instruments/Test Kits	34.51	35	65	74
17	Machinery & Industrial Goods	General Industrial Goods	35.91	34	60	77
18	Vehicles and Parts	Automotive Parts	50.62	19	49	92
19	Retailers	Food/Drug Retailers	22.73	76	118	41
20	Biopharmaceuticals	Small Molecule Therapeutics	25.72	67	104	51
21	Biopharmaceuticals	Gene Therapy	30.58	47	68	73
22	Software	Security	15.12	115	336	7
23	Retailers	Vehicle Parts Retailers/Vehicle Dealers	49.57	23	45	100
24	Travel and Leisure	Travel Arrangement/Tourism	17.53	100	179	25
25	Consumer Information Services	Entertainment	15.85	110	264	16
26	Financial Institutions & Services	Institutional Investment Services	19.79	91	129	37
27	Vehicles & Parts	Automobiles	157.79	2	30	130
28	Business Support Services	Procurement/Supply Chain	16.15	108	160	30
29	Media and Content	Broadcasting	39.56	28	37	112
30	Electronics & Computer Hardware	Electronic Components/Devices	15.08	117	194	23

Note: Rank here is based on the sum of both orders.

2.3.3. Gene Therapy: CRISPR, Cas9, and CAR-T

In the technology designated as “clustered, regularly interspaced, short palindromic repeats” (CRISPR) and the CRISPR-associated protein 9 (Cas9), the Cas9 enzyme functions as a fundamental part of a larger construct in which an RNA molecule guides the targeting of any possible matching DNA sequence. It is actually used to specify the critical site of cleavage. Because CRISPR–Cas has emerged as a highly flexible research tool for genome editing with the potential to enable researchers to manipulate the genome precisely, including the medical use of the system for directly treating genetic disorders, it has been publicized widely over fundamental parts of the CRISPR–Cas9 system [44, 45, 46].

The combination of chimeric antigen receptors (CARs) and artificial T cell receptors, CAR-T, is engineered receptors which graft an arbitrary specificity onto an immune effector cell (T cell). Typically, these receptors are used to graft the specificity of a monoclonal antibody onto a T cell, with transfer of their coding sequence facilitated by retroviral vectors. These receptors are chimeric because they comprise parts from different sources. The general premise of CAR T-Cells is rapid generation of T-Cells targeted to specific tumor cells. Once the T-Cell has been engineered to become a CAR T-Cell, it acquires supraphysiologic properties and develops the capability to act as a ‘Living Drug’ [47, 48, 49].

2.3.4. Immunotherapy and vaccines: Zika

Zika fever, also known as Zika virus disease or simply Zika, is an infectious disease caused by the Zika virus https://en.wikipedia.org/wiki/Zika_fever [50]. Symptoms include red eyes, joint pain, headache, fever, and a maculopapular rash [51, 52]. Although it has caused no associated fatalities https://en.wikipedia.org/wiki/Zika_fever [53],

mother-to-child transmission during pregnancy can cause microcephaly and other brain malformations in babies https://en.wikipedia.org/wiki/Zika_fever [54]. An outbreak that started in Brazil in 2015 spread to the Americas, Pacific, Asia, and Africa. This eventuality led to the World Health Organization’s declaration of Zika as a Public Health Emergency of International Concern in February 2016 [55]. Zika virus was little studied until the major outbreak. No specific antiviral treatment is available today [56].

A-4. Creation of author citation networks and Co-authorship networks related to keyword incidence in research papers relative to active financing in biopharmaceutical industry fields

Papers published during 2013–2017 including the previously described highest-growth keywords relative to actively financed biopharmaceutical industry fields of “exosome,” “microbiome,” “CRISPR,” “Cas9,” “CAR-T,” or “Zika” in the title, abstract, or keywords were extracted from the Web of Science. We targeted those papers for compilation of datasets to extract names of all authors and paper citation-related information to create author citation networks and co-authorship networks.

First, from the extracted data, we created lists of pairs (edge lists) of cited-and-citing papers. Using author lists for the respective papers, we constructed edge lists of cited-and-citing authors including all co-authors, irrespective of order, for all pairs of cited-and-citing papers. For instance, in a case in which paper A with five authors is cited by paper B with three authors, we created 15 (5×3) pairs as edge lists to build up a whole new author citation network based on them. For duplicate authors both in

Table 3. VentureSource keywords appearing twice or more for startups related to actively financing biopharmaceutical industry fields in 2017.

Appearance Frequency	Keywords (Appearance Frequency)
Pharmaceuticals	
5 or Greater	drug* (20), 'medic*' (12), 'cancer*' (12), 'therap*' (8), 'pharma*' (8), 'health*' (8), 'pharm*' (8), 'nan' (7), 'pain*' (5), 'vitamin*' (5), 'tumor*' (5), 'marijuana*' (5), 'cannabis*' (5), 'disease*' (5), 'pharmaceutic*' (5), 'protein*' (5)
3-4	vaccine* (4), 'supplement*' (4), 'skin*' (4), 'oncolog*' (4), 'antibiotic*' (4), 'nutrition*' (4), 'capsule*' (4), 'intermediate*' (4), 'biotech' (3), 'rare disease*' (3), 'bacteria*' (3), 'cardi*' (3), 'diabetes' (3), 'plant*' (3), 'treatment*' (3), 'oncology*' (3), 'therapeutic*' (3), 'child*' (3), 'cannabis' (3), 'medicine*' (3), 'treat*' (3), 'API*' (3), 'enzyme*' (3), 'immun*' (3), 'blood*' (3), 'tablet*' (3)
2	osteo*, 'OA', 'herpes', 'probiotic*', 'intestin*', 'digest*', 'molecule*', 'small molecule*', 'chemother*', 'immunoth*', 'opioid*', 'analges*', 'neuro*', 'molecul*', 'brain*', '*gum*', '*candy*', '*infect*', 'gynecolog*', 'clinic*', 'antimicrob*', 'bacteri*', 'broad-spectrum', 'cardio*', 'develop*', 'allerg*', 'neurolog*', 'cancer', 'respiratory*', 'asthma*', 'immune*', 'antibod*', 'obes*', 'compound*', 'antibiot*', 'bacteria', 'super*bug*', 'injection*', 'prescription', 'health', 'autoimmune*', 'Chinese medicine', 'herb*', 'diagnos*', 'exosome*', 'cosme*', 'tissue*', 'wrinkle*', 'API', 'glaucoma*', 'infect*', 'nutraceutic*'
Small Molecule Therapeutics	
5 or Greater	drug* (23), 'cancer*' (22), 'therap*' (20), 'molecule*' (15), 'disease*' (9), 'immun*' (8), 'neuro*' (8), 'health*' (8), 'cancer' (8), 'small molecule*' (7), 'tumor*' (6), 'treat*' (6), 'protein*' (6), 'antibod*' (6), 'cell*' (5), 'infect*' (5), 'molecul*' (5), 'inflam*' (5), 'treatment*' (5), 'cardio*' (5), 'small molecule' (5), 'oncolog*' (5)
3-4	therapeutic* (4), 'medic*' (4), 'pathogen*' (4), 'chronic*' (4), 'Alzheimer*' (4), 'biotech*' (3), 'immuno*' (3), 'ion channel' (3), 'inhibit*' (3), 'nano*' (3), 'immune*' (3), 'C5a' (3), 'terminal*' (3), '*infect*' (3), 'therapeut*' (3), 'bio*' (3), 'kinase' (3), 'tumor' (3), '*onco*' (3), 'diagnos*' (3), 'PSVT' (3), 'calcium*' (3), '*ventricul*' (3), 'brain*' (3), 'neurodegenerat*' (3)
2	biopharma*, 'affinity purification', 'bioprocess*', 'onco*', 'AI', 'substitut*', 'plant*', 'milk', 'cheese*', 'mayonnaise*', 'yogurt*', 'emul*', 'mitochondrial', 'eosinophil', 'leukocyte', '*immun*', 'skin*', 'spray*', 'resist*', 'bacteria*', 'antibiotic', 'AF', 'heart*', 'peptide*', '*cell*', 'derma*', 'ner*', 'molecul*', 'oncolog*', 'rare disease*', 'ossific*', 'fibrodysplasia', ' (RAR)', 'integrin', 'antibiotic*', 'breast*', '*cancer*', 'treatment', 'neurolog*', 'protein', 'lung', 'apoptosis', 'respiratory*', 'therapeutic', '*cancer*', 'fibrosis', 'GSK-3B', 'Glycogen Synthase Kinase', 'bipolar disorder', 'diabetes'
Biotechnology Therapeutics	
5 or Greater	cancer* (28), 'therap*' (20), 'drug*' (16), 'disease*' (12), 'biotech*' (11), 'cell*' (9), 'health*' (8), 'diseas*' (7), 'disorder*' (7), 'diagnos*' (7), 'gene*' (7), 'tumor*' (7), 'medic*' (6), 'bacteria*' (6), 'neuro*' (6), 'DNA' (6), 'oncolog*' (6), 'Alzheimer*' (5), 'cleantech' (5), 'treat*' (5), 'cancer' (5), 'tissue*' (5)
3-4	molecul* (4), 'immune*' (4), 'therapeut*' (4), 'antibod*' (4), 'therapeutic*' (4), 'industry focused products and services' (4), 'infect*' (4), 'regenerat*' (4), 'protein*' (4), 'microbiome' (4), 'treatment*' (4), 'immun*' (4), 'neurolog*' (3), 'Alzheimer*' (3), 'spinal cord*' (3), 'injur*' (3), '*amyloid*' (3), 'receptor' (3), 'central nervous system*' (3), 'CNS' (3), 'antibiotic*' (3), 'compound*' (3), 'onco*' (3), 'CMBC' (3), 'pharma*' (3), 'skin*' (3), 'molecul*' (3), 'vaccine*' (3), 'inhibitor*' (3), 'inflam*' (3), 'respirat*' (3), '*cancer*' (3), 'bone*' (3), 'pharm*' (3)
2	blood*, 'platelet*', 'manufact*', 'contract*', 'life scienc*', 'metaboli*', 'inflammat*', 'microbe*', 'inhal*', 'cardiovascular', 'PAF', 'biotech* atria*', 'arrhythmia', 'paediatr*', 'nephrolog*', 'renal*', 'neurologic*', 'orphan*', 'anaesthes*', 'tubulo*', 'oil', 'protein', 'insect*', 'agriculture', 'lysom*', 'stor*', 'HSP', 'misfold*', 'degenerat*', 'hear*', 'cochlear*', 'ear*', 'noise', 'restor*', 'cardiovascular*', 'myelofibrosis', 'MF', 'JAK2', 'ocular*', 'probiotic*', 'research*', 'clinical trial*', 'osteoporos*', 'hypoparathy*', 'hypoparathyroidism*', 'biopharm*', 'drug discover*', 'genomic*', '*microbiome*', 'th + F5eepuetic*', 'urea cycle disorder*', 'pathogenic*', '*bacteria*', '*health*', 'fish*', '*inflammation + F23*', 'detect*', 'vir*', 'T-cell*', '*immune*', 'dermatolog*', 'aesthetics', 'plasmotic', 'acne', 'hair removal', 'topical', 'vascular', 'COPD', 'addict*', 'opiate*', 'alcohol', 'abuse*', 'silk*', 'aesthetic*', 'defect*', 'diabetes', 'obes*', 'molecul*', 'rare disease*', '*skeletal*', 'drug', 'molecul*', 'metabolic', 'intestin*', 'antibody'
Gene Therapy	
5 or Greater	gene* (29), 'therap*' (19), 'cancer*' (16), 'DNA' (10), 'drug*' (7), 'genomic*' (5), 'cancer' (5), 'health*' (5), 'oncolog*' (5), 'cell*' (5)
3-4	DNA* (4), 'genom*' (4), 'patient*' (4), 'immun*' (4), 'CRISPR' (4), 'gene therap*' (4), 'cure' (4), 'therapy' (4), 'gen*' (4), 'disease*' (4), 'biotech*' (4), 'virus' (3), 'gene' (3), 'AAV' (3), 'virus*' (3), 'diseas*' (3), 'immune*' (3), 'brain' (3)
2	Parkinson*, 'DTC genetic test*', 'therapeutics*', 'genetic research*', 'medic*', 'adeno*', 'treat*', 'Cas9', 'duchenne', 'dystroph*', '8muscular disease*', 'viral*', 'tumor*', 'research*', 'animal*', 'livestock', 'agricultur*', 'breed*', '*medic*', 'retina*', 'dystrop*', 'choroid*', 'degenerat*', 'CHM', 'REP-1', 'protein*', 'rare disorder*', 'treatment*', 'adeno', 'medical', 'treatment', 'health', 'stem cell*', 'CAR-T', 'HIV', '*gene*', 'chimer*', 'receptor*', '*cell*', 'CAR', 'life scien*', 'genetic engineer*', 'personal* medic*', 'molec* bio*', 'HCP', 'cell protein', 'genome*', 'glioblastoma'
Immunotherapy/Vaccines	
5 or Greater	'cancer*' (31), 'immun*' (22), 'tumor*' (15), 'therap*' (14), 'antibod*' (13), 'disease*' (13), 'vaccine*' (12), 'infect*' (12), 'cancer' (11), 'drug*' (11), 'vaccin*' (10), 'virus*' (8), '*immun*' (8), 'antigen*' (7), 'T cell*' (6), 'cell*' (6), '*therap*' (5), 'protein*' (5)
3-4	'oncolog*' (4), 'allerg*' (4), '*cancer*' (4), 'immuno*' (4), 'disease' (3), 'biotech*' (3), 'inflammat*' (3), 'viral*' (3), '*tumor*' (3), 'antibody*' (3), 'ADC' (3), 'medicine*' (3), 'ag*' (3), 'immunosenescence' (3), '*infect*' (3), 'antibiotic*' (3), 'inflam*' (3), 'medic*' (3), 'prevent*' (3), 'RSV' (3), 'target*' (3), 'T-cell' (3)
2	'therapeutic*', 'oncology', 'RNA', 'DNA', 'HIV', 'TME', 'patient', 'pathogen*', 'bacteria*', 'monoclonal', 'complement system', 'COPD', 'AMD', 'PNH', 'virus', 'treatment', 'patho*', '*oncolog*', 'respirat*', 'licens*', 'rhinovirus', 'cold*', 'asthma*', '*vir*', '*mosquito*', 'Zika', 'Dengue*', '*fever*', 'Hepatit*', 'nanomedicine*', 'colorectal*', 'purif*', 'oncol*', 'viral', 'dendritic', 'biopharma*', 'diseas*', 'oncobi*', 'microb*', 'tumor', 'antigen'

cited and citing papers, we eliminated pairs of duplicates to produce pairs comprising different authors only. The method described above is applied to all pairs of cited-and-citing papers.

Subsequently, we introduced author networks of twokinds in an unweighted and undirected manner as described below. To observe how central the startup participant authors are and how they are distributed in the citation networks, we constructed new author citation networks with authors as nodes and with citation relations as edges (links) of these networks. The networks were based on edge lists of cited-and-citing authors with respect to each targeted keyword. At the same time, we constructed co-authorship networks from the papers above. The co-authorship network is a social network in which the authors, through participation in one or more publications through an indirect path, have been mutually linked, whereas author citation networks are based on citation relations among the authors. Therefore, we inferred that we

might observe different characteristics related to how central the startup participant authors are and how they are distributed, between author citation networks and co-authorship networks (Table 5).

A-5. Detecting startup participants in networks relative to emerging research topics in research papers related to actively financed biopharmaceutical industry fields

Using the VentureSource database again, we queried all names of nodes (authors) in the author citation networks and co-authorship networks to detect startup participants relative to emerging research topics represented by highest-growth keywords. We assumed that virtually any researcher with high centrality in those networks who participated in their startups were present in VentureSource because biopharmaceutical startups require large amounts of professional capital to conduct

Table 4. Keyword frequency growth multiple in web of science core collection related to active financing in biopharmaceutical industry fields during 2014–2017.

Rank	Keyword	Start Year Count	End Year Count	Growth (×)
Pharmaceuticals				
1	therap	1	6	6.00
2	cardi	2	7	3.50
3	super*bug	9	29	3.22
4	exosome	400	880	2.20
5	cosme	4	8	2.00
6	immun	7	14	2.00
7	allerg	1	2	2.00
8	antibod	1	2	2.00
9	obes	23	44	1.91
10	marijuana	993	1431	1.44
Small Molecule Therapeutics				
1	therap	1	6	6.00
2	inflam	1	3	3.00
3	immun	7	14	2.00
4	ventricul	1	2	2.00
5	PSVT	9	14	1.56
6	mayonnaise	41	60	1.46
7	yogurt	396	576	1.45
8	ossific	7	10	1.43
9	onco	126	178	1.41
10	rare disease	8249	11429	1.39
Biotechnology Therapeutics				
1	therap	1	6	6.00
2	inflam	1	3	3.00
3	microbiome	1732	4403	2.54
4	immun	7	14	2.00
5	injur	1	2	2.00
6	inflammat	1	2	2.00
7	diseas	4	8	2.00
8	obes	23	44	1.91
9	restor	23	35	1.52
10	paediatr	8	12	1.50
Gene Therapy				
1	therap	1	6	6.00
2	Cas9	448	2384	5.32
3	CAR-T	97	485	5.00
4	CRISPR	694	3138	4.52
5	DTC genetic test	13	29	2.23
6	immun	7	14	2.00
7	diseas	4	8	2.00
8	CHM	104	152	1.46
9	AAV	566	813	1.44
10	molec* bio	1335	1829	1.37
Immunotherapy/Vaccines				
1	Zika	26	2310	88.85
2	therap	1	6	6.00
3	inflam	1	3	3.00
4	allerg	1	2	2.00
5	immun	7	14	2.00
6	antibod	1	2	2.00
7	diseas	4	8	2.00
8	inflammat	1	2	2.00
9	TME	192	365	1.90
10	Dengue	1775	2729	1.54

Table 5. Comparison among research paper citation networks, author citation networks, and Co-authorship networks relative to growing keywords in actively financed biopharmaceutical industry fields in 2014–2017.

	Paper Citation Network	Author Citation Network	Co-authorship Network
<i>Exosome</i>			
Node Count	1,941	11,084	11,059
Edge Count	7,625	379,180	57,697
<i>Microbiome</i>			
Node Count	8,814	37,116	36,877
Edge Count	38,134	1,694,176	233,184
<i>CRISPR</i>			
Node Count	5,451	25,411	25,251
Edge Count	52,945	1,742,614	171,281
<i>Cas9</i>			
Node Count	3,974	19,893	19,808
Edge Count	38,856	1,415,583	133,925
<i>CAR-T</i>			
Node Count	685	3,302	3,281
Edge Count	4,377	277,377	25,106
<i>Zika</i>			
Node Count	2,987	13,137	12,943
Edge Count	29,196	1,864,665	102,358

substantive research and development activities. Dow Jones, the compiler of this database, had covered such industry-focused investors intensively and comprehensively.

The rankings of the startup participant authors were inferred from their degree centralities both in their author citation networks and co-authorship networks for each keyword showing increasing frequency, based on the sum of both regularized orders squared. Rankings of startup participant authors for each most growing keyword are shown in the Appendix.

3. Results

B-1. Constructing startup participant authors' distribution regarding degree centrality rankings in author citation networks and Co-authorship networks relative to growing keywords

For each growing keyword above, we mapped a scatter diagram of the distribution of startup participant authors in terms of their rankings of degree centralities both in the author citation networks and the co-authorship networks (Figure 2). We presented top authors alongside their degree centrality for the author citation networks and the co-authorship networks for each keyword (Table 6).

For startup activities in which these startup participant authors are engaged, one can extract information from the VentureSource database related to the role of participants, company overview, financing to date, and so on. Although we did not conduct thorough case studies across all startup companies engaged by all participants for the analyses described in this paper, we listed the 18 top-degree centrality startup participants with their 15 startups in each emerging research topic in Table 7 to verify the collectiveness and relevance of the startup participant author pool and the significance of the selected names by exemplifying several top startup participant authors for each emerging research topic. These startups have been successful either at raising venture capital, achieving an IPO, or being acquired by big pharmaceutical companies according to the VentureSource database. We can access up-to-date information for each relevant startup using VentureSource.

Results demonstrated that the startup participants are concentrated near the top of both networks, especially around the top 10%, as presented in Figure 2. Correlation between both the ranks of the author citation networks and those of the co-authorship networks to the same startup participants was to some degree: their correlation coefficients

were 0.337 (exosome), 0.464 (microbiome), 0.371 (CRISPR), 0.337 (Cas9), 0.505 (CAR-T), and 0.528 (Zika).

B-2. Hypothesis testing of top 10% degree centrality authors in author citation networks and Co-authorship networks related to "startup readiness" for each emerging research topic

From observations presented in B-1 regarding the six emerging research topics in the biopharmaceutical domain, we hypothesized that the proportion of startup participants is higher among authors of the top 10% degree centrality in both networks (designated hereinafter as "Dual Top 10% Authors") than it is among authors who have no such high centrality.

To conduct testing of our hypothesis, we used Fisher's exact test to infer significance of differences in the observed proportions. Fisher's exact test, a test of statistical significance used for analysis of contingency tables, assesses significance of deviation from a null hypothesis, or P -value, calculated exactly as long as the contingency tables' row and column totals are fixed, rather than relying on an approximation, as does a chi-square approximation [57, 58, 59]. We calculated the probability P that the number of startup participants is equal to or exceeds the observed number among "Dual Top 10% Authors," with the null hypothesis that startup participants are equally likely to be distributed among authors in both networks irrespective of their degree centralities. Additionally, we calculated how much higher the odds of being a startup participant is among "Dual Top 10% Authors" compared with other authors, i.e., odds ratio [60, 61] too.

Following are findings for the six emerging research topics in Table 8. We infer that the results we observed in their odds ratios were statistically significant.

- (i) The P -values were 1.875e-08 (exosome), 4.41e-12 (microbiome), 2.2e-16 (CRISPR), 2.2e-16 (Cas9), 0.000519 (CAR-T), and 0.0382 (Zika), all of which were equal to or less than three decimal places except for Zika's P -value, which remained less than 0.05, the number used as the cutoff in most statistical hypothesis testing.
- (ii) Odds ratios across all the emerging research topics were 3.166 (exosome), 2.214 (microbiome), 5.360 (CRISPR), 5.880 (Cas9), 2.450 (CAR-T) and 1.675 (Zika), all of which observed a higher

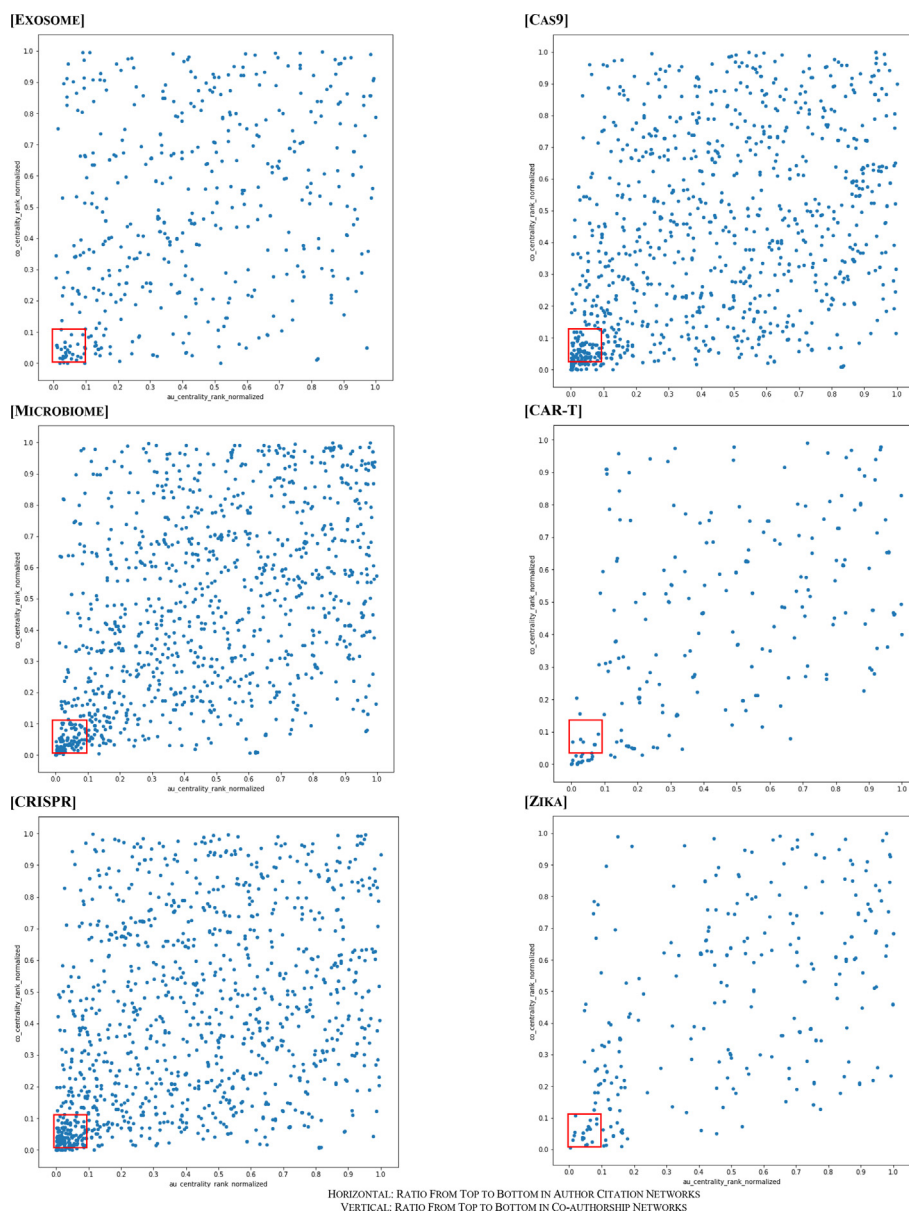


Figure 2. Scatter Diagram, Distribution of Startup Participant Authors' Degree Centrality in Author Citation Networks & Co-authorship Networks based on Ratio from Top to Bottom, in Emerging Research Topics in Actively Financed Biopharmaceutical Industry Fields in 2014–2017

startup participant ratio in “Dual Top 10% Authors” than in other authors.

4. Discussion

Our hypothesis was that researchers who are academically central in both author citation networks and co-authorship networks tend to have higher startup readiness in such research topics that appearance frequency is emerging while belonging to actively financed biopharmaceutical industrial segments. Compared to those who are not central, they tend to participate in startups, transforming their research outcomes into commercial applications in emerging fundamental research fields, particularly in fields such as life sciences. Life sciences are user-inspired. They tend to transform scientific research outcomes into practical usage. They attract venture capital financing, as shown in Table 2.

Earlier studies conducted for detection of emerging research fields have used bibliometric approaches, such as the work of Shibata et al. (2008), who used a topological clustering method to divide citation networks into clusters and who then tracked the positions of papers in

each cluster [62]. Shibata et al. (2011) calculated network centralities, so-called betweenness centralities, of papers with respect to regenerative medicine [63]. Sasaki et al. (2016) calculated nine kinds of network centrality for papers related to photovoltaic solar cells: degree centrality, betweenness centrality, closeness centrality, eigenvector centrality, network constraint, clustering coefficient, Page rank, hub score, and authority score [64]. Although these studies did not address each researcher's preparedness to create startups, to facilitate decision-making processes they proposed prediction models to identify emerging promising studies that might attract many citations. The biopharmaceutical domain is characterized by intense scientific activity and entrepreneurship, as described in the Introduction. Therefore, we strove to emphasize the application of such network centrality and to develop paper-related individual factors to assess academic researchers' startup readiness [4]. Results suggest that scientists with top degree centrality in author citation networks and co-authorship networks have a higher startup participant ratio than other scientists in these specifically selected biopharmaceutical domains (Table 8).

Table 6. Top authors' degree centrality in author citation networks & Co-authorship networks in emerging research topics in actively financed biopharmaceutical industry fields for 2014–2017.

Author Citation Network			Co-Authorship Network		
Rank	Author	DegCent	Rank	Author	DegCent
[Exosome]					
1	Ochiya, Takahiro	0.1115	1	Jensen, Torben Heick	0.0109
2	Minn, Andy J.	0.0992	2	Ochiya, Takahiro	0.0088
3	Ghajar, Cyrus M.	0.0809	3	Liu, Yutao	0.0077
4	Mallya, Kavita	0.0809	4	Minn, Andy J.	0.0065
5	Zhang, Tuo	0.0809	5	Lotvall, Jan	0.0061
6	Zhang, Haiying	0.0809	6	Sandelin, Albin	0.0060
e	Pantel, Klaus	0.0809	7	Chen, Yun	0.0059
8	Muller, Volkmar	0.0809	8	Helwa, Inas	0.0057
9	Fodstad, Oystein	0.0809	9	Gho, Yong Song	0.0056
10	Hernandez, Jonathan	0.0809	10	Dismuke, W. Michael	0.0055
[Microbiome]					
1	Knight, Rob	0.2343	1	Knight, Rob	0.0207
2	Huttenhower, Curtis	0.1531	2	Huttenhower, Curtis	0.0134
3	Gevers, Dirk	0.1248	3	Wang, Jun	0.0096
4	Xavier, Ramnik J.	0.1242	4	Gilbert, Jack A.	0.0089
5	Fischbach, Michael A.	0.1121	5	Kristiansen, Karsten	0.0085
e	Turnbaugh, Peter J.	0.1047	6	Bork, Peer	0.0082
7	Wang, Jun	0.1028	7	Xavier, Ramnik J.	0.0082
8	Gootenberg, David B.	0.1007	8	Raes, Jeroen	0.0077
9	Gonzalez, Antonio	0.1005	9	Dore, Joel	0.0072
10	David, Lawrence A.	0.0999	10	Li, Junhua	0.0069
[CRISPR]					
1	Zhang, Feng	0.4556	1	Zhang, Feng	0.0146
e	Doudna, Jennifer A.	0.2834	2	Katsanis, Nicholas	0.0122
e	Lander, Eric S.	0.2649	3	Kim, Jin-Soo	0.0097
4	Scott, David A.	0.2538	4	Zhang, Wei	0.0082
e	Joung, J. Keith	0.2525	5	Huang, Xingxu	0.0076
6	Hsu, Patrick D.	0.2455	6	Davis, Erica E.	0.0076
7	Shalem, Ophir	0.2422	7	Li, Wei	0.0069
8	Kim, Jin-Soo	0.2179	8	Root, David E.	0.0069
9	Sanjana, Neville E.	0.2037	9	Doudna, Jennifer A.	0.0067
10	Charpentier, Emmanuelle	0.2024	10	Wang, Yan	0.0064
[Cas9]					
1	Zhang, Feng	0.4728	1	Zhang, Feng	0.0163
e	Doudna, Jennifer A.	0.2942	2	Katsanis, Nicholas	0.0126
3	Scott, David A.	0.2922	3	Kim, Jin-Soo	0.0111
4	Hsu, Patrick D.	0.2851	4	Huang, Xingxu	0.0098
5	Joung, J. Keith	0.2786	5	Davis, Erica E.	0.0097
6	Lander, Eric S.	0.2762	6	Zhang, Wei	0.0091
7	Kim, Jin-Soo	0.2495	7	Wang, Yan	0.0081
8	Shalem, Ophir	0.2464	8	Cho, Megan T.	0.0081
9	Sharp, Phillip A.	0.2251	9	Hildebrandt, Friedhelm	0.0079
10	Root, David E.	0.2214	10	Liu, Wei	0.0074
[CAR-T]					
1	June, Carl H.	0.6319	1	June, Carl H.	0.0790
2	Kochenderfer, James N.	0.5780	2	Lacey, Simon F.	0.0473
3	Sadelain, Michel	0.5532	3	Dotti, Gianpietro	0.0430
4	Stetler-Stevenson, Maryalice	0.5153	4	Cooper, Laurence J. N.	0.0418
5	Rosenberg, Steven A.	0.5089	5	Jensen, Michael C.	0.0415
6	Mackall, Crystal L.	0.4998	6	Levine, Bruce L.	0.0381
7	Fry, Terry J.	0.4995	7	Sadelain, Michel	0.0372
8	Davila, Marco L.	0.4647	8	Brenner, Malcolm K.	0.0348
e	Riviere, Isabelle	0.4553	9	Kochenderfer, James N.	0.0341
10	Feldman, Steven A.	0.4553	10	Maus, Marcela V.	0.0323
[Zika]					
1	Musso, Didier	0.5221	1	Honein, Margaret A.	0.0240

(continued on next page)

Table 6 (continued)

Author Citation Network			Co-Authorship Network		
Rank	Author	DegCent	Rank	Author	DegCent
2	Cao-Lormeau, Van-Mai	0.4769	2	Jamieson, Denise J.	0.0237
3	Jamieson, Denise J.	0.4133	3	Meaney-Delman, Dana	0.0226
4	Honein, Margaret A.	0.4024	4	Shi, Pei-Yong	0.0172
5	Teissier, Anita	0.3835	5	Kraemer, Moritz U. G.	0.0158
6	Roche, Claudine	0.3795	6	Diamond, Michael S.	0.0155
7	Petersen, Lyle R.	0.3757	7	Khan, Kamran	0.0153
8	Mallet, Henri-Pierre	0.3442	8	Moore, Cynthia A.	0.0152
9	Diamond, Michael S.	0.3409	9	Weaver, Scott C.	0.0151
10	Vasilakis, Nikos	0.3264	10	Oduyabo, Titilope	0.0149

As presented in Figure 2, which portrays startup participant authors' degree centrality positions in the author citation networks and in co-authorship networks related to scientific topics emerging in the biopharmaceutical domain, we observed that both networks have a common tendency by which higher degree centrality is associated with higher density of startup participant authors: higher "startup readiness." Table 9 presents correlation coefficients between the degree centrality of the author citation networks and that of the co-authorship networks to the same authors: 0.434 (exosome), 0.628 (microbiome), 0.349 (CRISPR), 0.322 (Cas9), 0.619 (CAR-T), and 0.556 (Zika). All of those except for CRISPR and Cas9 were higher than those of the degree centralities' ranks, as discussed in B-1. It might also be of additional value to discuss the potential use of patents as features to assess startup readiness as well. As discussed earlier in the Introduction, because it is commonly observed that in biotechnology papers linked to a patent receive more citations than those without a patent link [5], we can assume that authors' patenting activities are correlated with such authors' standing in the networks described previously. In fact, as Table 9 shows, CRISPR and Cas9 had higher correlation coefficients between authors' degree centrality in the author citation networks and the number of patents such authors invented: 0.413 (CRISPR) and 0.399 (Cas9), than those between authors' degree centrality of the author citation networks and that of the co-authorship networks: 0.349 (CRISPR) and 0.322 (Cas9). In other research topics, however, correlation coefficients between authors' degree centrality in the author citation networks and the number of patents such authors invented (0.069 (exosome), 0.019 (microbiome), 0.138 (CAR-T), and 0.078 (Zika)), were much lower than those between the authors' degree centrality of the author citation networks and that of the co-authorship networks: 0.434 (exosome), 0.628 (microbiome), 0.619 (CAR-T), and 0.556 (Zika).

Given such stark difference in correlation coefficients between CRISPR, Cas9, and the other research topics, we eventually did not implement patent-related factors across them. As shown in Table 8, CRISPR and Cas9 particularly were the top two research topics related to startup participant ratio in "Dual Top 10% Authors" compared to other authors.

However, results also show that correlation between the ranks of author citation networks and those of co-authorship networks to the same startup participants were not strong. This result might derive from the fundamental difference between author citation networks that we constructed from paper citation networks and co-authorship networks. Because the author citation networks were created based on citations among papers, linkages among the authors in the networks were open among authors who have mutual interest, irrespective of their human relations or organizational connections. By contrast, the co-authorship networks were based mainly on human relations and organizations to which the authors belong. They are expected to be more limited and closed networks than the author citation networks. These results suggest that both networks are useful complementarily when we seek authors in emerging research fields who might be startup participants. This complementarity apparently contributed to our Fisher's exact test results

and to calculation of the odds ratio, which distinguished "Dual Top 10% Authors" from authors who were not so centrally important.

Results of our Fisher's exact test for each emerging research topic, or each keyword with rapidly growing usage in the biopharmaceutical domain support our hypothesis that "Dual Top 10% Authors" have higher "startup-readiness." We infer that all results are significant. On a side note, with respect to "Zika," whose P -value 0.0382 was way larger than those of the other emerging research topics, all of which were equal to or less than three decimal places. We presume that "Zika," as the name of a specific viral disease, differs from the other keywords "exosome," "microbiome," "CRISPR," "Cas9," and "CAR-T," all of which can be useful for therapy or well-being. We infer that the P -value of "Zika" is high because of the great number of research projects examining "Zika" that might not fit any immediate use for society.

In addition to the points presented above, we acknowledge that variables other than centrality in networks of academic author citation and co-authorship might be used. Moreover, they might affect startup readiness in general. However, our approach can work well specifically for startups in commercialization-oriented yet fundamentally scientific research domains such as the biopharmaceutical domain, as exemplified herein.

This method can extract the latest names of research topics and researchers comprehensively, but easily in a scalable computational manner. By contrast, earlier methods have collected data using conventional methods such as personal interviews, reading published papers, or conducting field projects in person, offering only limited instantaneity, comprehensiveness, and scalability.

We proposed a computational method that uses openly available databases of startup finances and academic research papers combined, to explore the "startup readiness" of emerging research topics and researchers in user-inspired fundamental scientific research such as life sciences. It is suitable to assess aspects of academic entrepreneurship enables us to boost the collection of real-time data related to relevant science and industry segments and researcher pools, both in greater detail and on a larger scale.

5. Conclusions

To seek emerging scientific research topics for which startup activities are active, we presented a method based on keyword analyses and academic databases to sort scientific keywords associated with active financing activities and academic research activities. First, we constructed cited-and-citing author networks and co-authorship networks based on paper citation networks relative to the biopharmaceutical domain. Then we analyzed the degree centralities of nodes of the biopharmaceutical-related startup participant authors in their author citation networks as well as co-authorship networks to detect signals of startup readiness in emerging research domains.

Results demonstrated that the startup participant authors exhibited high centrality for author citation networks and co-authorship networks related to biopharmaceuticals, especially within the top 10% of centrality for both. Furthermore, P -values obtained from our Fisher's exact test

Table 7. Startup Activities for which Top Startup Participants Are Engaged Relative to Each Emerging Research Topic.

Startup Participant	Role	Startup Company	Company Overview Brief Description	Most Recent Financing (MM)
Exosome				
Zhang, Bin	VP	Cisen Pharmaceutical Co. Ltd.	Manufacturer of chemical pharmaceutical agents and related products such as non- polyvinyl chloride (PVC) soft bag infusions, plastic bottle infusions, lyophilized powder injections, tablets, ointments, eye drops, and capsules	09/29/2017 IPO 1166.00RMB (Chinese Yuan)
Chen, Wei	Board Member, Outsider	Immunophotonics, Inc.	Developer of a cancer treatment	09/04/2014 VC 1st 2.49 USD
Xu, Bin	EVP	Grandhope Biotech Co. Ltd.	Provider of medical materials and devices for the treatment of damaged tissue and organs	07/06/2011 IPO 278.4 RMB (Chinese Yuan)
Johansson, Henrik J.	Unknown Executive	Halo Genomics AB	Developer of targeted re-sequencing technology for DNA sequencing	Acquired by Agilent
Microbiome				
Xavier, Ramnik J.	Cofounder	Jnana Therapeutics	Developer of drug that target cellular proteins	12/14/2017 VC 1st 50.00 USD
de Vos, Willem M.	Chairman, Scientific Advisory Board	MicroDish BV	Developer of micro-engineered culture chips and nanoscale reagents aiming to improve microbial culture	03/31/2011 VC 1 st N.A.
Mazmanian, Sarkis	Director	AxialBiotherapeutics Inc.	Developer of biotherapeutic that target neurological diseases and disorders	06/22/2017 VC 1st 19.20 USD
CRISPR				
Zhang, Feng	Cofounder	Editas Medicine	Developer of human therapeutics based on genome editing technologies	02/03/2016 IPO 94.40 USD
Doudna, Jennifer A.	Cofounder	Editas Medicine	Same as above	Same as above
	Cofounder	Intellia Therapeutics, Inc.	Provider of CRISPR-Cas9 focused biotechnology	05/06/2016 IPO 108.00USD
Joung, J. Keith	Cofounder	Editas Medicine	Developer of human therapeutics based on genome editing technologies	Same as above
Cas9				
Zhang, Feng	Cofounder	Editas Medicine	Developer of human therapeutics based on genome editing technologies	Same as above
Doudna, Jennifer A.	Cofounder	Editas Medicine	Same as above	Same as above
	Cofounder	Intellia Therapeutics, Inc.	Provider of CRISPR-Cas9 focused biotechnology	Same as above
Joung, J. Keith	Cofounder	Editas Medicine	Developer of human therapeutics based on genome editing technologies	Same as above
CAR-T				
June, Carl H.	Cofounder	Tmunity Therapeutics	Developer of T-cell immunotherapies	01/23/2018 VC 2 nd 100.00USD
Sadelain, Michel	Cofounder	Juno Therapeutics Inc.	Developer of medicines to treat cancer	12/19/2014 IPO 264.55USD
Riviere, Isabelle	Cofounder	Juno Therapeutics Inc.	Same as above	Same as above
Zika				
Osorio, Jorge E.	Chief Scientific Officer	Inviragen Inc.	Developer of vaccines to protect against infectious diseases	06/07/2013 Acquired by Takeda
Schinazi, Raymond F.	Founder	Pharmasset Inc.	Developer of drugs to treat viral infections	04/27/2007 IPO 45.00 USD
	Board Member, Outsider	Gliknik Inc.	Developer of cancer, autoimmune, and inflammatory disease therapies	09/30/2013 Corporate Partnership 25.00
	Board Member, Outsider	ReViral Ltd.	Developer of antiviral drugs focused on diseases caused by the respiratory syncytial virus (RSV)	09/08/2015 VC 1 st 21.00 USD

supported our hypothesis: the proportion of startup participants is higher for authors in the top 10% of centrality of both networks than for other authors. Therefore, to detect researchers with “startup readiness” in an emerging research topic, one can assess their combined degree-centrality in author citation networks and co-authorship networks. Because correlation between both the ranks of the author citation networks and those

of the co-authorship networks is not high with respect to degree centrality to the same startup participants, complementarity between these two networks can strengthen the collectiveness and relevancy of the author pool as startup participant candidates.

This analytical scheme is unique: Our citation network constructs author citation networks derived from paper citation networks.

Table 8. Contingency tables related to the number of startup participants and non-participants for dual top 10% authors and others with P-value and odds ratio for each research topic.

	Startup Participant	Non-Participant
<i>Exosome</i>		
Dual Top 10% Authors	37	262
Other Authors	439	9,845
P-value: 1.875e-08	Odds ratio: 3.166	
<i>Microbiome</i>		
Dual Top 10% Authors	107	1605
Other Authors	995	33045
P-value: 4.41e-12	Odds ratio: 2.214	
<i>CRISPR</i>		
Dual Top 10% Authors	142	611
Other Authors	1018	23480
P-value: 2.2e-16	Odds ratio: 5.360	
<i>Cas9</i>		
Dual Top 10% Authors	118	402
Other Authors	870	17430
P-value: 2.2e-16	Odds ratio: 5.880	
<i>CAR-T</i>		
Dual Top 10% Authors	23	123
Other Authors	206	2700
P-value: 0.000519	Odds ratio: 2.450	
<i>Zika</i>		
Dual Top 10% Authors	20	550
Other Authors	257	11839
P-value: 0.0382	Odds ratio: 1.675	

Table 9. Correlation coefficients of authors' degree centrality in author citation networks, degree centrality in Co-authorship networks & number of patents they invented, in emerging research topics in actively financed biopharmaceutical industry fields in 2014–2017.

	Citation Deg. Cent.	Coauth. Deg. Cent.	Number of Patents
<i>[Exosome]</i>			
Citation Deg. Cent.	1		
Coauth. Deg. Cent.	0.434	1	
Number of Patents	0.069	0.057	1
<i>[Microbiome]</i>			
Citation Deg. Cent.	1		
Coauth. Deg. Cent.	0.628	1	
Number of Patents	0.019	0.014	1
<i>[CRISPR]</i>			
Citation Deg. Cent.	1		
Coauth. Deg. Cent.	0.349	1	
Number of Patents	0.413	0.213	1
<i>[Cas9]</i>			
Citation Deg. Cent.	1		
Coauth. Deg. Cent.	0.322	1	
Number of Patents	0.399	0.204	1
<i>[CAR-T]</i>			
Citation Deg. Cent.	1		
Coauth. Deg. Cent.	0.619	1	
Number of Patents	0.138	0.257	1
<i>[Zika]</i>			
Citation Deg. Cent.	1		
Coauth. Deg. Cent.	0.556	1	
Number of Patents	0.078	0.064	1

Moreover, our analyses specifically examine authors, not papers. The method calculates all authors' degree centralities exhaustively and can thereby reveal startup readiness. Using author names rather than published papers can make this method scalable into various fields and make

it adaptable to situations in which the publication cycle is short and the publications are numerous. In contrast to other variables representing the maturity/level of technology such as TRL and researcher reputation, which have been tacit, difficult to attain or difficult to measure, we use

variables such as financing activities, research topics, and researchers' degree centralities. For that reason, and because our approach is based on publicly available digital data, people with little or no knowledge related to specific disciplines, industries, and regions can use this method to detect research topics and researchers with startup readiness. Furthermore, our computational approach might update necessary information in real time, except for the time lag until publication, although such a time lag is decreasing by virtue of the increasing popularity of open journals. Our method, which is useful for researchers and practitioners participating in research-based startups, can contribute to a rich, convenient, and dynamic understanding of academic entrepreneurship.

Because our case study specifically examined biopharmaceutical scientists in certain research topics, the findings here might not be generalizable across all biopharmaceutical research topics, let alone across all disciplines. However, we selected the scientists by objective evaluation based upon ranking, regarding the degree to which financing activities are active in their industry segments, the degree to which academic attention is growing in their research topics, and how high their degree centrality is in the respective networks of author citation and co-authorship from their paper citation networks. Our approach might be generalizable or highly suggestive to other similarly user-inspired fundamental scientific research fields with abundant potential for commercialization and academic entrepreneurship. Moreover, even though this paper examined scientists with specifically selected research topics, our method could be applicable to other biopharmaceutical research topics as well as those whose academic attention is similarly emerging while belonging to actively invested biopharmaceutical fields.

Potential disparities among paper author numbers across research topics and our selection of industry segments demand further study because they can influence results. The collectiveness and the relevancy of the author pool and the significance of the selected names requires further research from perspectives such as their startups' completeness, suitability and importance. Strengthening the legitimacy of our methods requires further research to address circumstances under which papers have numerous or limited authors, to widen our methods' applicability to other industries or research topics, to ascertain the effects of selection and influential factors, and to survey the whole pools of relevant authors and startups at levels ranging from network structure to commercial, financial viability/success.

Declarations

Author contribution statement

T. Goji: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Y. Hayashi: Performed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.

I. Sakata: Conceived and designed the experiments; Analyzed and interpreted the data; Wrote the paper.

Funding statement

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Competing interest statement

The authors declare no conflict of interest.

Additional information

Supplementary content related to this article has been published online at <https://doi.org/10.1016/j.heliyon.2020.e04160>.

Acknowledgements

We thank Dr. Atsushi Usami (The University of Tokyo Edge Capital Partners, Co., Ltd. (UTECC)) for his insights on CRISPR, Cas9 and microbiome, Hiroko Yamano (Innovation Policy Research Center, School of Engineering, The University of Tokyo) for her support of conceptualization, and Takanari Matsuda (Department of Technology Management for Innovation, School of Engineering, The University of Tokyo) for his support of conceptualization and software analysis.

References

- [1] J.C. Mankins, Technology readiness assessments: a retrospective, *Acta Astronaut.* 65 (9–10) (2009) 1216–1223.
- [2] H. Nakamura, Y. Kajikawa, S. Suzuki, Multi-level perspectives with technology readiness measures for aviation innovation, *Sustainability Science* 8 (1) (2013) 87–101.
- [3] T.E. Stuart, W.W. Ding, When do scientists become entrepreneurs? The social structural antecedents of commercial activity in the academic life sciences, *Am. J. Sociol.* 112 (1) (2006) 97–144.
- [4] T. Goji, T. Matsuda, I. Sakata, Measuring "start-up readiness" of scientific research-based start-ups using analysis of citation networks: case study of CRISPR-Cas9, in: *Proceedings of PICMET '17: Technology Management for Interconnected World*, 2017.
- [5] T. Magerman, B. Van Looy, K. Debackere, Does involvement in patenting jeopardize one's academic footprint? An analysis of patent-paper pairs in biotechnology, *Res. Pol.* 44 (9) (2015) 1702–1713.
- [6] J. Barney, Firm resources and sustained competitive advantage, *J. Manag.* 17 (1) (1991) 99–120.
- [7] B. Kogut, U. Zander, Knowledge of the firm, combinative capabilities, and the replication of technology, *Organ. Sci.* 3 (3) (1992) 383–397.
- [8] K.R. Conner, C.K. Prahalad, A resource-based theory of the firm: knowledge versus opportunism, *Organ. Sci.* 7 (5) (1996) 477–501.
- [9] R.M. Grant, Toward a knowledge-based theory of the firm, *Strat. Manag. J.* 17 (S2) (1996) 109–122.
- [10] R. Landry, N. Amara, I. Rherrad, Why are some university researchers more likely to create spin-offs than others? Evidence from Canadian universities, *Res. Pol.* 35 (10) (2006) 1599–1615.
- [11] M. Knockaert, A. Spithoven, B. Clarysse, The knowledge paradox explored: what is impeding the creation of ICT spin-offs? *Technol. Anal. Strat. Manag.* 22 (4) (2010) 479–493.
- [12] E. Rasmussen, O.J. Borch, University capabilities in facilitating entrepreneurship: a longitudinal study of spin-off ventures at mid-range universities, *Res. Pol.* 39 (5) (2010) 602–612.
- [13] T. Huynh, D. Patton, D. Arias-Aranda, L.M. Molina-Fernández, University spin-off's performance: capabilities and networks of founding teams at creation phase, *J. Bus. Res.* 78 (2017) 10–22.
- [14] C. Corsi, A. Prencipe, M.J. Rodríguez-Gulías, D. Rodeiro-Pazos, S. Fernández-López, Growth of KIBS and non-KIBS firms: evidence from university spin-offs, *Serv. Ind. J.* 39 (1) (2019) 43–64.
- [15] F.T. Rothaermel, S.D. Agung, L. Jiang, University entrepreneurship: a taxonomy of the literature, *Ind. Corp. Change* 16 (4) (2007) 691–791.
- [16] J. Bercovitz, M. Feldman, Academic entrepreneurs: organizational change at the individual level, *Organ. Sci.* 19 (1) (2008) 69–89.
- [17] S. Jain, G. George, M. Maltarich, Academics or entrepreneurs? Investigating role identity modification of university scientists involved in commercialization activity, *Res. Pol.* 38 (6) (2009) 922–935.
- [18] B. Clarysse, V. Tartari, A. Salter, The impact of entrepreneurial capacity, experience and organizational support on academic entrepreneurship, *Res. Pol.* 40 (8) (2011) 1084–1093.
- [19] M. Abreu, V. Grinevich, The nature of academic entrepreneurship in the UK: widening the specifically examine entrepreneurial activities, *Res. Pol.* 42 (2) (2013) 408–422.
- [20] T.T. Aldridge, D. Audretsch, S. Desai, V. Nadella, Scientist entrepreneurship across scientific fields, *J. Technol. Tran.* 39 (6) (2014) 819–835.
- [21] R.A. Rader, (Re) defining biopharmaceuticals, *Nat. Biotechnol.* 26 (7) (2008) 743.
- [22] P.E. Stephan, S. Gurmu, A.J. Sumell, G. Black, Who's patenting in the university? Evidence from the survey of doctorate recipients, *Econ. Innovat. N. Technol.* 16 (2) (2007) 71–99.
- [23] D.E. Stokes, *Pasteur's Quadrant: Basic Science and Technological Innovation*, Brookings Institution Press, Washington, DC, 1997.
- [24] R. Henderson, A.B. Jaffe, M. Trajtenberg, Universities as a source of commercial technology: a detailed analysis of university patenting, 1965–1988, *Rev. Econ. Stat.* 80 (1) (1998) 119–127.
- [25] A.N. Link, D.S. Siegel, B. Bozeman, An empirical analysis of the propensity of academics to engage in informal university technology transfer, *Ind. Corp. Change* 16 (4) (2007) 641–655.
- [26] W.W. Powell, J. Owen-Smith, Universities and the market for intellectual property in the life sciences, *J. Pol. Anal. Manag.* 17 (2) (1998) 253–277.
- [27] Report from "Council for Promoting Ventures Innovating Healthcare" in July 2016, Ministry of Health, Labor and Welfare (MHLW) of Japan, 2016.

- [28] L.C. Freeman, Centrality in social networks conceptual clarification, *Soc. Network.* 1 (3) (1979) 215–239.
- [29] L.C. Freeman, A set of measures of centrality based on betweenness, *Sociometry* 40 (1) (1977) 35–41.
- [30] P. Bonacich, Technique for analyzing overlapping memberships, *Socio. Methodol.* 4 (1972) 176–185.
- [31] R.S. Burt, Structural holes and good ideas, *Am. J. Sociol.* 110 (2) (2004) 349–399.
- [32] D.J. Watts, S.H. Strogatz, Collective dynamics of ‘small-world’ networks, *Nature* 393 (1998) 440–442.
- [33] Y. Ding, Applying weighted PageRank to author citation networks, *J. Am. Soc. Inf. Sci. Technol.* 62 (2) (2011) 236–245.
- [34] G. Cainelli, M.A. Maggioni, T.E. Uberti, The strength of strong ties: how co-authorship affects productivity of academic economists, *Scientometrics* 102 (2015) 673–699.
- [35] E. van der Pol, A.N. Böing, P. Harrison, A. Sturk, R. Nieuwland, Classification, functions, and clinical relevance of extracellular vesicles, *Pharmacol. Rev.* 64 (3) (2012) 676–705.
- [36] S. Keller, M.P. Sanderson, A. Stoek, P. Altevogt, Exosomes: from biogenesis and secretion to biological function, *Immunol. Lett.* 107 (2) (2006) 102–108.
- [37] A.M. Booth, Y. Fang, J.K. Fallon, J.M. Yang, J.E. Hildreth, S.J. Gould, Exosomes and HIV Gag bud from endosome-like domains of the T cell plasma membrane, *J. Cell Biol.* 172 (6) (2006) 923–935.
- [38] J. Lederberg, A.T. McCray, Ome SweetOmic – a genealogical treasury of words, *Scientist* 15 (7) (2001) 8.
- [39] J. Peterson, S. Garges, M. Giovanni, P. McInnes, L. Wang, J.A. Schloss, C.C. Baker, The NIH human microbiome project, *Genome Res.* 19 (12) (2009) 2317–2323.
- [40] F. Backhed, R.E. Ley, J.L. Sonnenburg, D.A. Peterson, J.I. Gordon, Host-bacterial mutualism in the human intestine, *Science* 307 (5717) (2005) 1915–1920.
- [41] P.J. Turnbaugh, R.E. Ley, M. Hamady, C.M. Fraser-Liggett, R. Knight, J.I. Gordon, The human microbiome project, *Nature* 449 (7164) (2007) 804.
- [42] R.E. Ley, D.A. Peterson, J.I. Gordon, Ecological and evolutionary forces shaping microbial diversity in the human intestine, *Cell* 124 (4) (2006) 837–848.
- [43] T. Nakatsuji, T.H. Chen, S. Narala, K.A. Chun, T. Yun, F. Shafiq, J.N. Kim, Antimicrobials from human skin commensal bacteria protect against *Staphylococcus aureus* and are deficient in atopic dermatitis, *Sci. Transl. Med.* 9 (378) (2017), eaah4680.
- [44] K.J. Egelie, G.D. Graff, S.P. Strand, B. Johansen, The emerging patent landscape of CRISPR-Cas gene editing technology, *Nat. Biotechnol.* 34 (10) (2016) 1025–1031.
- [45] E. Deltcheva, K. Chylinski, C.M. Sharma, K. Gonzales, Y. Chao, Z.A. Pirzada, E. Charpentier, CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III, *Nature* 471 (7340) (2011) 602.
- [46] L. Cong, F.A. Ran, D. Cox, S. Lin, R. Barretto, N. Habib, F. Zhang, Multiplex genome engineering using CRISPR/Cas systems, *Science* 339 (6121) (2013) 819–823.
- [47] M. Sadelain, R. Brentjens, I. Riviere, The basic principles of chimeric antigen receptor (CAR) design, *Canc. Discov.* 3 (4) (2013) 388–398.
- [48] S. Srivastava, S.R. Riddell, Engineering CAR-T cells: design concepts, *Trends Immunol.* 36 (8) (2015) 494–502.
- [49] J. Hartmann, M. Schüßler-Lenz, A. Bondanza, C.J. Buchholz, Clinical development of CAR T cells – challenges and opportunities in translating innovative treatment concepts, *EMBO Mol. Med.* 9 (9) (2017) 1183–1197.
- [50] Zika Situation Report on 5 February 2016, World Health Organization, 2016.
- [51] L.H. Chen, D.H. Hamer, Zika virus: rapid spread in the western hemisphere, *Ann. Intern. Med.* 164 (2016) 613.
- [52] D. Musso, E.J. Nilles, V.-M. Cao-Lorreau, Rapid spread of emerging Zika virus in the Pacific area, *Clin. Microbiol. Infect.* 20 (10) (2014) O595–O596.
- [53] Factsheet about Zika Virus Disease, European Centre for Disease Prevention and Control, 2016 updated on 23 June 2016.
- [54] P. Brasil, J.P. Pereira Jr., C. Raja-Gabaglia, et al., Zika virus infection in pregnant women in Rio de Janeiro – preliminary report, *N. Engl. J. Med.* 375 (2016) 2321–2334.
- [55] WHO Statement on the First Meeting of the International Health Regulations Emergency Committee on Zika Virus, World Health Organization, 1 February 2016.
- [56] V. Sikka, V.K. Chattu, R.K. Popli, et al., The emergence of Zika virus as a global health security threat: a review and a consensus statement of the INDUSEM Joint Working Group (JWG), *J. Global Infect. Dis.* 8 (1) (2016) 3–15.
- [57] R.A. Fisher, On the interpretation of χ^2 from contingency tables, and the calculation of P. J. Roy, *Stat. Soc.* 85 (1) (1922) 87–94.
- [58] R.A. Fisher, *Statistical Methods for Research Workers*, Genesis Publishing Pvt. Ltd, 1925.
- [59] A. Agresti, A survey of exact inference for contingency tables, *Stat. Sci.* 7 (1) (1992) 131–153.
- [60] F. Mosteller, Association and estimation in contingency tables, *J. Am. Stat. Assoc.* 63 (321) (1968) 1–28.
- [61] A.W.F. Edwards, The measure of association in a 2×2 table, *J. Roy. Stat. Soc.* 126 (1) (1963) 109–114.
- [62] N. Shibata, Y. Kajikawa, Y. Takeda, K. Matsushima, Detecting emerging research fronts based on topological measures in citation networks of scientific publications, *Technovation* 28 (11) (2008) 758–775.
- [63] N. Shibata, Y. Kajikawa, Y. Takeda, I. Sakata, K. Matsushima, Detecting emerging research fronts in regenerative medicine by citation network analysis of scientific publications, in: PICMET’09-2009 Portland International Conference on Management of Engineering & Technology, IEEE, 2009, August, pp. 2964–2976.
- [64] H. Sasaki, T. Hara, I. Sakata, Identifying emerging research related to solar cells field using a machine learning approach, *J. Sustain. Development of Energy, Water and Environ. Sys.* 4 (4) (2016) 418–429.