Renewable and Sustainable Energy Reviews 77 (2017) 318–326

Contents lists available at ScienceDirect

Renewable and Sustainable Energy Reviews

journal homepage: www.elsevier.com/locate/rser

The emergence of patent races in lignocellulosic biofuels, 2002–2015

Hannes Toivanen[a,b,c,*], Michael Novotny[d]

[a] VTT Technical Research Centre of Finland, P.O. Box 1000, 02044-VTT, Finland
[b] Lappeenranta University of Technology, School of Business and Strategy, Lappeenranta, Finland
[c] Teqmine Analytics Oy, Helsinki, Finland
[d] KTH Royal Institute of Technology, Department of Industrial Economics & Management, Stockholm, Sweden

ARTICLE INFO

Keywords:
Biofuels
Lignocellulose
Patent analysis
Patent landscape
Policy
Emerging technology
Business strategy

ABSTRACT

How does increasing economic and technological interest in biofuels shape the nature of the intellectual property rights (IPR) in the industry? Is the technological nature of biofuel patents and inventions, as well as the business itself undergoing a transformation? This article provides a patent analysis of lignocellulosic biofuels with U.S. patent publications between 2002 and 2015 in order to shed light on the broader economic and regulatory factors affecting the development of new technologies in the area. Patent applications in the technology have increased about eightfold in this period and count about 130–150 per year currently, and could soon reach 200 annual filings. Specifically, we analyse in what ways the nature of lignocellulosic biofuel technologies is changing, and our results suggest that this business is indeed being transformed by increasing research and development (R & D) and IPR efforts, material in an evident patent race. We document a relatively small, but nascent technology, with some key technology areas increasing between four- and tenfold, over the last decade. Technologically leading countries are the U.S., followed by Germany, Japan, France and the U.K. We argue that intensified global and industry-wide claims for IPR reveal an ongoing patent race with multiple implications for the industry and engineering community. Most importantly, industry's technological interdependence is likely to increase as the likelihood for broad, exclusive patent regimes diminishes, making the industry more likely to explore increasingly collaborative technological solutions when carrying out R & D and investing in new production facilities in the future.

1. Introduction

In global energy research, renewable energy sources have since 2000 become the fastest growing nexus of technological research and invention, evident in rapidly growing patent rates that outpace other energy technologies [1]. This global search for economically, politically and environmentally sustainable energy sources underpins broad global changes, and the nature and direction of innovation in renewable biofuels is, therefore, of great importance for policy makers and industry leaders as well as the research and development community. Increased patenting has raised the issue of how emerging intellectual property rights (IPR) regimes may affect the rate, direction, and diffusion of innovation. One manifestation of this has been the policy debate on IPR rights in renewable energy technologies. This discussion has been polarized on to what extent IPR rights work as incentives for firms to invest in renewable biofuel technologies, and to what extent patents effectively deter critical innovation [2,3].

In this context, this paper offers a careful patent analysis of the lignocellulosic biofuels technology. Lignocellulose-based biofuels embody one of the most potent new sources of renewable energy. Abundant in nature, they are one of the largest sources of carbohydrates, yet the exploitation of lignocellulose on a large scale and in an economically feasible fashion is limited by technological difficulties of efficiently converting biomass into ethanol and other chemicals. The potential of lignocellulosic biofuels as energy and a bioeconomy platform technology, as well as the increasing possibilities of mobilizing related production technologies on an industrial scale, have fuelled and continue to fuel an increasing technological and commercial interest in the area [4,5].

With patent analysis, this article documents a rapidly increasing technological effort in the area, whereby technology can best be characterized as "emerging technology", and explores how this intensified activity and interest is reshaping the broader social, economic and regulatory environment of lignocellulosic biofuels. By using advanced patent analysis tools, such as a mix of supervised and unsupervised machine learning techniques and text-mining, on the complete United States Patent and Trademark Office (USPTO) patent database consisting of approximately 7 million full-text patents, we

* Corresponding author at: Teqmine Analytics Oy, Helsinki, Finland.
E-mail addresses: Hannes.Toivanen@vtt.fi (H. Toivanen), Novotny@kth.se (M. Novotny).

http://dx.doi.org/10.1016/j.rser.2017.03.089
Received 1 April 2016; Received in revised form 30 January 2017; Accepted 20 March 2017

describe a technological area characterized by rapidly increasing patenting rates and clear signs of patent races. Such a phenomenon signals increasing industrial technological viability, but also suggests changes in the nature of inventions and intellectual property rights claims. To summarize, inventive steps in the industry are becoming smaller, more incremental and more tightly connected, and the technology as a whole is becoming increasingly covered by exclusive IPR. If true, such changed circumstances are bound to limit the strategic choices available to firms and research organizations, and are likely to induce more collaborative practices between leading holders of IPRs.

2. Background

2.1. Wood-based biofuels and biorefineries

Recent interest in lignocellulosic biofuels is part of a global search for technologically, economically, and environmentally viable energy sources. One manifestation driving increased energy demand has been the wide changes in agricultural practices, especially irrigation systems [6]. As a non-food crop, wood-based renewable biofuels embody a particularly attractive solution, and were already identified as such at the beginning of the chemical wood processing industry. The first non-food, crop-based, biofuel solutions were introduced roughly 100 years ago, when bioethanol was one of the major industrial products from wood-based, sulphite and pulp mills in Scandinavia, Germany and the USA [7].

After World War II, the success of the oil and petroleum industries, as well as the pulp and paper industry completely shifted to sulphate pulp processes [8], which quickly eclipsed substantial efforts to develop and expand wood-based biofuel technologies [9]. Continued interest in bioethanol based on other crops secured the development of key technologies, mostly fermentation technologies, the key nexus being the Brazilian effort to develop sugarcane-based bioethanol.

The rise of second generation biofuel technologies, fuelled by the global search for more environmentally sound fuel sources, especially those focused on the exploitation of various biomasses, has more recently expanded the technological search and business interest in lignocellulosic biofuels [10]. Basically, any lignocellulosic biomass – agricultural residues, municipal waste, wood residues, (micro)algae, switchgrass, to mention the most common feedstocks – can be extracted and converted into carbohydrate sugars, alcohols, lignin phenols and oils in a first stage and biofuels in a second stage [7].

More recently, the food/fuel rivalry has triggered a scramble for alternative feedstocks, i.e. lignocellulosic material, which form the basis for the next generations of biofuels. Lignocellulosic components, such as cellulose, hemicellulose and lignin, represent an interesting and challenging opportunity for producing sustainable and renewable liquid fuels.

Wood-based biofuels enjoy several potential advantages. One advantage is that carbohydrates (cellulose and hemicelluloses sugars) are the most abundant organic feedstock on Earth, with lignin, comprising about 25% of all biomass compounds, being the second most abundant. Roughly 75% of the lignocellulosic feedstock is composed of carbohydrates. Some technological bottlenecks arrest their exploitation, especially the difficulty of efficiently separating out the lignin, with only about five percent of all carbohydrates being used by humankind.

There are three main technology platforms for biofuels: (a) hydrolysis – enzymatic (biochemical) or thermochemical hydrolysis, (b) pyrolysis (solid biofuels or bio-oil), and (c) gasification [7,11]. Liquid biofuels are usually produced via gasification or hydrolysis technologies. Hydrolysis or fermentation is usually related to ethanol production, and gasification is a syngas conversion of carbon monoxide, carbon dioxide, methane and hydrogen. Syngas, as a gas, can serve as fuel for power or as a liquid fuel, in particular, methanol. Methanol is a versatile platform compound for several chemicals and fuels (e.g. acetic acid and dimethyl ether) (Table 1).

Dimethyl ether (DME) has in turn emerged as an efficient energy carrier and transportation fuel in the form of biodiesel [7,12]. The advantages of biodiesel fuels are the same as for ethanol. Ethanol can easily be used by the established transportation system, that is in fossil based liquid blends, for transportation fuels at established gasoline pumps. Furthermore, it can be processed from almost any abundant lignocellulosic material (e.g. wood and/or agricultural residues or energy crops such as algae). This, so called Third Generation of biofuels, is based on upgrades in the production of biomass, which takes advantage of specifically engineered energy crops such as algae [13].

Biofuel production is often part of biorefinery infrastructures in facilities that convert biomass feedstock, including lignocellulosic material such as wood, into a wide range of valuable and sustainable materials, chemicals and multiple fuels [7]. The biorefinery concept embraces several renewable raw materials, separation and conversion technologies, and intermediate and final products. A biorefinery potentially has the advantage of producing more classes of products than can petroleum refineries, thus offering more opportunities for product development and diversification. A bottleneck consists of the logistics of heterogeneous raw materials which makes access to biomass a strategic factor when locating a biorefinery production unit. Currently, there are not many global examples of cost-competitive alternatives to oil refineries, but in the 2000s, the uncertainties related to climate change, volatile energy prices and fossil-based energy supplies, ‘cotton peaks’ and the like, have contributed to making biomass appealing as a feedstock for many industries and industrial processes. In a lignocellulosic context, a range of technological barriers and economic and industrial opportunities have been identified, and resulted in a nexus of highly active research and development [14].

Given the political, economic and societal pressures to move towards cleaner and environmentally sustainable energy systems and economic factors, the business interest to develop a cost-efficient and technologically sound wood-based biorefinery concept has surged during the last decade. This interest has materialized as a rapidly growing research effort, as evidenced by the surge in scientific publications and patent applications. Naturally, when investment in a wood-based biomass biorefinery easily exceeds 1 billion euros, the need to secure a return on investments for the development of necessary new technologies affects the balance between scientific research and claim on exclusive intellectual property rights.

2.2. Emerging technologies and patent races

The notion of “emerging technologies” is usually applied in a practical sense to define a given technological nexus being capable of overturning or replacing existing mainstream solutions. It is invoked, typically, as a framework to guide expectations and focus R & D and business efforts. In the context of this paper, we label lignocellulosic biofuels as an “emerging technology” because of their potential as a new biofuel production method, as well as because of the rapidly increasing volume of research and development efforts (see below).

Whereas nanocellulose or biofuels are perhaps the current embodiments of emerging technologies in the forest based industries, Toivanen [8,9] has argued that the industry has, since the 1850s, been swept by a succession of emerging technologies. The first wave was the introduction of wood pulp and the first sulphite pulping process, followed by different product innovations (newsprint, tissue paper, corrugated board, etc.). This was followed by the second wave that included the sulphate pulping process, and eventually by the third wave, the introduction of engineered fast-growing tree species and adapted pulp processes. Each of these waves either substituted existing production processes or products or introduced completely new ones. Moreover, each of these emerging waves sparked an industry-wide

Table 1
Biorefinery categories, biochemical/biofuels production, their technologies, feedstock and some examples of applications.

Category	Technology	Feedstock	Biochemical	Application
Sugars/starch (1st & 2nd generation)	Hydrolysis (biochemical)	Carbohydrates (wheat, sugarbean, non-food crops such as sugarcane)	Ethanol, buthanol, lactic acid	Fuels, solvents, paints, etc
Oil/lipids (1st generation)	Thermochemical	Soybean, sunflower,	Resins, fatty acids, glycerine	Lubricants, pharma, coating, personal care; fuels
Forest biomass derivatives (3rd generation)	Thermochemical (hydrolysis/ gasification)	Woody lignocellulose (wood, black liquor, tall oil)	Turpentine, rosin, ethanol, lignin, methanol, cellulose, hemicellulose	Soaps, detergents, textiles, plastics, fertilizers, transportation fuels
Other biomass derivatives (3rd generation)	Thermochemical (Gasification/ hydrolysis)	Any lignocellulose (Algae, biomass residues, agricultural crop residues, switchgrass, bagasse)	Methanol, acetic acid, olefins, ethanol,	Transportation fuels, e.g. car/aviation fuel via syngas (gasification of hydrogen, methane, CO or CO2) plastics, etc

transition, manifest in organizational and regulatory change. Novotny and Laestadius have shown, with extensive qualitative analysis, that the industry's transition to convert biomass into energy and bioproducts is accompanied by many similar management, organizational and policy challenges [15].

In a biofuels context, several patent studies have confirmed increasing patenting rates and changing business dynamics. Biofuel and biohydrogen technologies increased substantially in the early 2000s, and the technology has significantly matured, prompting strengthened commercialization strategies [16]. Similar emerging and growth dynamics have been confirmed for biomass fermentation technologies [17], as well as in the broader and aggregate "clean energy" field [18]. These studies document an increasing patent rate in high-priority biofuel technology areas, often suggesting deeply changing business and industry dynamics.

Industry transitions [19] underpinned by novel technological departures, and the emergence of new, complex, technological, production systems, are typically accompanied by intense business competition. This is particularly true in cases requiring heavy capital investments so as to achieve sufficient economies of scale, and under these circumstances, large and small firms have been shown to follow distinct strategic paths in order to secure long-term advantages from their investments in research and development [20]. In the pulp and paper industry, such transitions engendered repeatedly intensive, litigious, and lengthy patent races [8].

Patent races should be interpreted as manifestations of firms' attempts to secure a competitive advantage over competitors, and cause a rapid increase in the number of patent applications. Often, such races reveal a deeper change. Industry leaders' intensified research and development efforts can also suggest the imminent commercial viability of specific production technologies. However, if no firm or patent manages to claim very broad exclusive IPR rights to the emerging technology and effectively negates competitors' efforts, increased research and patent activity is likely to impact the very nature of research and development activity, patent claims, and eventually the corporate strategy of leading firms [21].

In this paper, we argue that the nature of the lignocellulosic biofuels patent race is likely to create two specific outcomes. First, key IPR rights will be divided into a number of strategic IPR portfolios held by competing firms, and, secondly the building of major production systems will require some sort of cooperative arrangements between the owners of these portfolios. Furthermore, as no firm in the industry is likely to have a clear monopolistic grip on technological solutions, as a whole the industry is likely to explore increasingly collaborative arrangements in R & D and production technologies.

Finally, our paper explores the role of technological intelligence for the biofuel community, and demonstrates the advantages of versatile Big Data approaches to mapping the technological, IPR and business environment. When developing and choosing the best strategies for moving forward within the emerging biofuel space, research organizations, firms, and governments must obtain information about the broader technology and business environment, and in this regard, we also briefly outline the economic, strategy and policy implications.

3. Data and methods

3.1. Data

Patent data was obtained from the USPTO patent database maintained by Teqmine Analytics Ltd, which is based on USPTO-issued XML publications. The data for analysis was prepared in five steps for the final analysis. In the *first phase*, we text mined all USPTO full-text publications between 1990 and 2015 (inclusive) with the key words string "*lign*" and "*fuel*". This search yielded 82,580 publications. In the *second phase*, this data was processed with Latent Dirichlet Allocation (LDA, see below for full description) and all the records were allocated to 20 topics. By visual inspection, and using four reference patents (e.g. US20140256979A1 and US20130333652A1), we identified two topics capturing patents focussing on pyrolysis and catalysts in a biomass and hydrocarbon context. Patents in this area totalled 5144. In the *third phase*, the 4803 patents were divided with LDA into 10 topics (see below for methods).

The *fourth* phase consisted of selecting the final patent pool providing the best evidence on recent evolution of the lignocellulosic fuel technology. To this end, data was limited to publications between 2002 and 2015 (inclusive) and kind A1, B1, P1 and P2. These cover first-time publications of utility and plant patent applications, as well as first-time publications of granted patents without prior publications. Because the data continued to include noise and patents of no immediate relevance to lignocellulosic fuel processing technologies, we excluded redundant patents from the data.

To this end, we visually inspected the Topic Modelling classification, as well as analysing the classification of our above-mentioned reference, US patents. More importantly, we controlled our LDA results by comparing our results with the International Patent Classification (IPC) at record level. This work confirmed that Topic 9 effectively captured all patents that were of relevance to us, and, as verified in Table 2, depict topic distribution of patents by the World International Patent Organization, and introduce 33 aggregate technology categories from the IPC classifications (Table 2).

The fifth and final stage of data processing consisted of eliminating as much redundant data as possible from the selected Topic 9. Consisting of 2562 records, Topic 9 is effectively contaminated with unrelated patents addressing mostly fuel cell, pharmaceutical, and lubricant patents. While pharmaceutical and lubricant patent areas are technologically closely related to lignocellulosic fuel processing, fuel cell technologies carry only linguistic similarity. Elimination of such "noise" from the patent data is an enduring challenge in patent analysis and a prerequisite of credible analysis. Thus, we resorted to using a sample of record level IPC classifications to choose patents for the final analysis data.

Final lignocellulosic biofuel technology patent data consisted of 1069 records, in our judgement, included at most very few non-relevant patents and effectively captured relevant USPTO patent publications with a bearing on lignocellulosic fuel processing and production technologies.

3.2. Methods – patent analysis

Patent analysis of emerging technologies, or competitive technological intelligence, is a relatively well-established methodological tradition, though one that is currently impacted by the introduction of a range of new "Big Data" methodologies, such as natural language processing (NLP) and machine learning approaches. The methodological transition in patent analysis is essentially about completing patent analysis more efficiently, faster, and with a higher level of accuracy, all with the potential of improving the practical value of patent analysis for engineers and business leaders [22].

Whereas traditional patent analysis methods make extensive use of the meta data at the patent publication level, such as technology classes, keywords, and citation data, in order to map and classify data for analytical purposes, Big Data approaches rely heavily on unsuper-

Table 2
Comparing LDA results with WIPO IPC technology classes.
Source: USPTO, Teqmine Analytics Oy, Authors.

		Topic										
No	IPC Sub-Field	1	2	3	4	5	6	7	8	9	10	Total
1	Analysis of biological materials		8	4	11			25	44	24	2	118
2	Audio-visual technology					1						1
3	Basic materials chemistry	35	19	19		94		2	23	865	450	1507
4	Biotechnology		44	48	62	16	39	40	77	193	6	525
5	Chemical engineering		17			24			1	353		395
6	Civil engineering					2				50		52
7	Computer technology		12					2	25	5		44
8	Control		2			1			2			5
9	Electrical machinery, apparatus, energy		2			6				120		128
10	Engines, pumps, turbines		1							25		26
11	Environmental technology		9			2				69		80
12	Food chemistry		21	2		17			14	110	5	169
13	Furniture, games					6			1	1		8
14	Handling		3	1		3			1	10		18
15	IT methods for management		1						4	1		6
16	Machine tools		1			4			1	19		25
17	Macromolecular chemistry, polymers		9	2		118			1	496	2	628
18	Materials, metallurgy		5			13				138		156
19	Measurement		6	1		1		1	6	19		34
20	Mechanical elements								1	17		18
21	Medical technology		21	9		3		10	159	17	1	220
22	Micro-structural and nano-technology		5	6		6		1	3	21	1	43
24	Optics		1			2				20		23
25	Organic fine chemistry	23	30	154	30	21	23	12	46	645	232	1216
26	Other consumer goods			2		6			3	27		38
27	Other special machines	1	5			19		3	4	88	23	143
28	Pharmaceuticals	3	59	272	29	8	43	58	336	62	40	910
29	Semiconductors		1			3				12		16
30	Surface technology, coating		13			30			1	166	4	214
31	Telecommunications								1			1
32	Textile and paper machines		13			22				83		118
33	Thermal processes and apparatus					1				6		7
34	Transport					32				10		42
	Total	62	308	520	132	461	105	154	754	3672	766	

Note: Patents are counted once for each class they have received a IPC classification.

vised or semi-supervised machine learning approaches. The latter approaches offer several advantages. These include such things as the ability to rapidly process very large data, such as reading millions of full-text patent texts, which is not easily possible for humans, and often enabling the possibility to analyse technologies from a range of perspectives. Another advantage relates to improved and higher rates of semantic discovery, as NLP approaches analyse all data and enable the linking of data pools based on semantic similarities. In contrast, traditional patent analysis methods often suffer from the "street light" syndrome, where one searches only within the predefined area [23,24].

We employed a mix of unsupervised and supervised, i.e. semi-supervised learning methods in processing patent data, as described in our data sections. Our unsupervised learning methods rely on Latent Dirichlet Allocation (LDA), which is a topic model that identifies latent patterns from semantic text. Whereas a range of algorithms cover what are known as "Topic models", LDA is by far the most familiar of them [25]. LDA is a three-layer Bayesian model that is now widely used in different applications, such as text mining, bioinformatics, and image processing. We use LDA to pass the data through iterative analysis cycles, whereby, we eliminate redundant, and identify relevant, data. However, the relevant data on lignocellulosic biofuels is, in the end, so small, as well as linguistically overlapping with other fields, such as lubricants and pharmaceuticals, that in the final data processing stages, it is practical also to use meta-data approaches.

On the practical side, our use of a mix of unsupervised and supervised classification methods offer two important advantages. *First*, we were able to screen whole patent data to identify the technology area where potentially relevant inventions on lignocellulosic biofuels could be found, regardless of the type of technology classifications assigned to patents by examining offices or inventors. *Second*, we were able to discard non-relevant patent data from key technology classes, as several important classifications for lignocellulosic biofuels cover a very broad range of technologies, industries and biomass sources. Thus, our methodological approach also demonstrates the value and applicability of novel Big Data approaches for patent analysis in biofuel technologies.

4. Results

4.1. Emerging technology area

Lignocellulosic biofuels embody a relatively small but nascent technology area. Annual patenting in the technology increased about eightfold between 2002 and 2015 to about 130–160 annual filings. (Fig. 1.) An increase in patenting rate of this size can only be interpreted as significant, although we have not verified for divisional applications or patent families, and it suggests an intensified technological and economic interest in the technology. Given its recent growth and the likelihood of continued intense technological research and development, lignocellulosic biofuels are appropriately characterized as an emerging technology with a high probability of patent races.

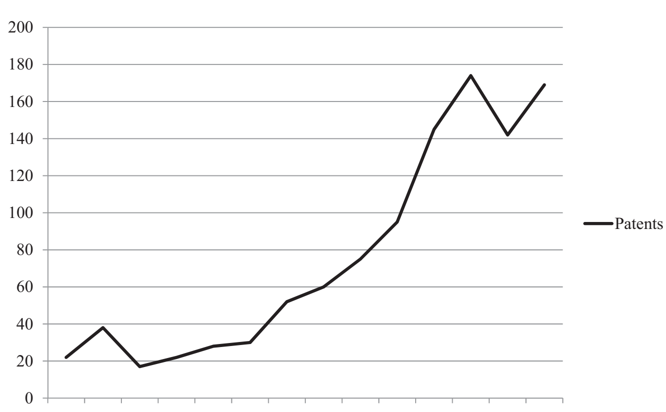

Fig. 1. Lignocellulosic biofuel patents, 2002–2015.
Source: USPTO, Teqmine Analytics Oy, Authors

While several key patents in lignocellulosic biofuels processing and production had already been filed in the 1990s, which often made wide-ranging claims for the key technological processes, our results demonstrate a continued search for new solutions and improvements for technological challenges, and suggest that industry expansion continues to be slowed down because of serious technical problems. Perhaps the best-known example of such problems is the experience of Kior's Mississippi plant and its eventual bankruptcy. While predicting a future patenting rate is relatively uncertain, a simple projection based on our results would suggest that patenting, in this area, is likely stay above 130 patents over the next few years; and could soon reach 200 annual filings. Annual patenting rates also attest to the intensification of economic interests in lignocellulosic biofuels, as well as the fact that patents are considered by industry players to be a key strategic asset.

Another matter is that an increased patenting rate could also be interpreted as a changed nature of technology or inventions, and that it could also carry implications for competitive strategies. Whereas early lignocellulosic biofuel patents often embodied wider claims and were filed within a relatively vacant patent space, contemporary patent applications must navigate an increasingly tense and occupied IPR space. This development can translate to increasingly narrow and focussed claims in patents, as well as establishing increasing interdependency between key patent portfolio owners and prompting the need for a range of new collaboration strategies.

While a detailed analysis of the business and IPR strategy implications of the significant growth of lignocellulosic biofuel patenting is beyond the scope of this paper, it is clear that such implications, as considered above, exist and weigh heavily on corporate decisions in the renewable biofuels industry.

4.2. Leading countries

Which countries are spearheading the emerging lignocellulosic biofuel technologies, and are these countries following specific technological trajectories? To this end, Table 3 depicts 20 leading inventor countries in 2002–2015.

U.S.-based inventors have contributed to about 60% of all lignocellulosic patents, making them the largest single inventor group by far. With USPTO patents, this is often self-evident, as U.S-based inventors are more likely to apply for a U.S. patent than are foreign-based inventors. The second largest inventor group is Germany-based, followed closely by Japan, France, and Great Britain-based inventors. The seven largest non-U.S. inventor countries produce about one quarter of all patent publications. Within the list, one is perhaps surprised that Sweden and Finland, traditional technology leaders in forest-based technologies, only rank as a tied 11th and 14th place, and are surpassed by Canada, the Netherlands, Belgium, Israel, and China. (Table 3).

Geographically, the sources of inventions in lignocellulosic biofuels are relatively diffuse, with very few true regional clusters. In the U.S., most patents are by inventors based in Houston (76 patents), followed by San Diego (49), Pasadena (39), Madison (38), Wilmington (23) and Atlanta (20). Leading cities outside the U.S. are, in order, Mannheim, Tokyo, Cambridge, Lyon, Amsterdam and Jerusalem, Heidelberg and Ludwigshafen, with each having produced between 10 and 14 patents. Importantly, patents from these cities are, as a rule, obtained as private corporations as assignees.

The role played by key assignees is suggestive too, although it is difficult to provide a reliable assessment of key assignees in the data, as USPTO practice allows filing of patent applications in a manner that conceals possible real assignees. However, the largest assignee in our data is the Celanese Corporation with 47 publications. Of these, 39

Table 3
Lignocellulosic biofuel patents by inventor country, 2002–2015. 20 leading inventor countries.
Source: USPTO, Teqmine Analytics Oy, Authors.

Country	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	Total
US	11	28	11	15	20	13	27	34	36	42	98	111	97	102	645
DE	2		1	1		2	9	4	8	6	9	10	6	11	69
JP	2	1		2	2	7	3	5	4	11	7	7	2	8	61
FR				1	1	4	4	2	5	12	4	9	6	3	51
GB			2	1	2	4	3	3	4	8	4	9	4	6	50
CA	1	1		1		2	4	3	3	1	5	7	4	4	36
NL		1		1	4				4	7	6	2	4	6	35
BE					1	1	2			2	8	7	4	9	34
IL							2	1	1	1	6	2	4	3	20
CN				2				1	4	4	3	2		1	17
BR							1	2	1	1	2	4	2		13
IT			1		1				1	1	1	5	1	2	13
SE	1			1				2	1			2	1	5	13
FI						1		1	1	2	1	2	4		12
IN		1						1	1	1	2	2	2	2	12
N		8	3												11
KR						1					1	2	3	3	10
SA								1				1	1	5	8
CH						1			1				1	3	6
ES							1		2	2	1				6
Total	22	38	17	22	28	30	52	60	75	95	145	174	142	169	

NOTE: Countries are represented by 2-letter standard abbreviations.

were published in a single year, suggesting a determined corporate tactic. Celanese is followed by BASF (22 patents), Shell (20), UOP (18) and IPF (13). Of these, only BASF has had a long history of patenting in the area, as the other firms have filed almost all of their patents in the last five years, within the time frame covered in this analysis. The increasing corporate interest in lignocellulosic biofuel patenting is strongly suggestive of patent race dynamics.

Given the fact that R & D and patenting in lignocellulosic biofuels is increasing substantially, several key countries and cities are, in practice, being transformed into repeat invention hubs, producing an increasing number of intellectual property claims within the business area. This interpretation is strengthened by the fact that many of the leading actors in key regions are firms, either incumbent industry leaders or highly specialized new entrants. This development is likely to have a significant impact on the nature of the lignocellulosic biofuel business, as well as shaping the nature of patenting and technological trajectories.

4.3. Inventive frontiers in lignocellulosic biofuels

Lignocellulosic biofuel technologies are evidently growing in size, to the extent that they are best described as an “emerging technology”. Here, we turn to analysing whether this growth is accompanied by a shifting technological focus, and, more specifically, determining what are the key technological focus areas within lignocellulosic biofuels R & D efforts, and whether they are displaying different dynamics? To answer this question, Table 4 shows annual filings by key IPC class level for the largest technology groups.

Lignocellulosic biofuel patenting concentrates on four technology class areas, of which the C07C “Acyclic or carbocyclic compounds”, especially covering electrolysis approaches, is the most significant, covering almost 40% of all lignocellulosic biofuel patents. Almost non-existent as lignocellulosic biofuel technology in the early 2000s, contemporary patenting in the IPC class is between 60 and 100, and this demonstrates how rapidly technological and economic interest in the area has proliferated. This class and the patents selected through our methods constitute a key technology in lignocellulosic biofuels.

The “carbocyclic” part of the IPC code “C07C” covers the aromatic ring of carbon atoms found in many biomolecules that chemically may substitute major petroleum-based molecules, such as the hydrocarbons benzene and cyclopropane, which make up a large part (almost 50%) of gasoline composition, particularly in high octane gasoline. Products from this category may be used as e.g. bioethanol or blended with gasoline. The technological and business importance of this IPC technology class is underscored by the factor by which patent filings in it have surged in recent years. (Table 4).

A second key area is established by the IPC class area C10L, “Fuels not otherwise provided for”, which covers some 28% of the total patenting in the lignocellulosic biofuels area, and totals some 25–40 patents per year. The importance of this technology class for the biofuels technology and business community is also reflected in patenting activity. Although this technology area has substantially increased since 2002, it cannot be characterized as “emerging”. The area's patenting has declined substantially in the early 2000s, and has only recently begun to recover. While we have made an effort to eliminate pure fuel additives and lubricants from the selection, this group continues, to some degree, to include patents that focus more on these technology areas rather than biofuels. However, it is often difficult to establish a clear demarcation between wood-based fuels, fuel additives and lubricants, as they might be “by-products” of the same or of a closely related production system.

The IPC class B01J, “Chemical or physical processes, e.g. Catalysis…” embodies another clearly emerging, yet small, technology area. Patenting in the area has increased within lignocellulosic biofuels from practically nothing to about 20–40 per year, and it captures more than 10% of the whole patent group. This technology class is well described as “nascent”, as it did not exist in lignocellulosic biofuels in the early 2000s, but ranks as one of the four largest ones today.

The C12P class, “Fermentation or enzyme-using processes to synthetize….” also captures some 10% of the total lignocellulosic patenting volume. It is best described as moderately emerging, increasing from a few annual patents to about 1 to −30 between 2002 and 2015.

Enzymatic processes and enzymatic catalysis, related to IPC classes B01J and C12P, are used for two reasons. They are first of all part of one of the most promising ecological pre-treatment processes for fermentation technologies, e.g. bioethanol technologies. Secondly, with enzymatic treatment, the purpose is to obtain higher fuel yields – or in general fermentable sugars – at low energy costs. In particular, the emerging technology, the booming biomass-based fuels (bioethanol

Table 4
Lignocellulosic biofuel patents by key IPC classes, 2002–2015.
Source: USPTO, Teqmine Analytics Oy, Authors.

IPC	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	Total	%
C07C	1	2	2	2	9	6	10	16	29	36	65	99	63	91	431	40,32%
C10L	13	27	13	7	10	7	20	25	21	26	37	32	23	43	304	28,44%
B01J					1	2	1	6	3	10	20	42	15	27	127	11,88%
C12P	2	1		5	2	3	8	6	11	13	27	14	14	21	127	11,88%
C08F	4	2		6	3	9	9	4	7	17	16	8	7	13	105	9,82%
C10G		1		2	3	1	3	7	3	8	12	17	16	28	101	9,45%
C07D						2	1	2	4	8	15	10	17	21	80	7,48%
C13K		2			1		1	3	6	6	15	6	13	12	65	6,08%
C08B			1				2	4	7	4	9	12	10	4	53	4,96%
C02F	1	3	1		2	3	1	1	2	3	9	6	11	8	51	4,77%
C08G						1		1	4	3	6	7	5	7	34	3,18%
C08L					1	1	2	1	1	5	2	4	4	8	29	2,71%
C10M					1	1	3	2	1	3	9	3	1	4	28	2,62%
C07H						1	1	4	1	1	13	1	1	2	25	2,34%
B01D					2		2		1	2	4	3	2	6	22	2,06%
C09K							1	4	2	4	3	2	2	3	21	1,96%
C11B						1		3	1		6	5	2	3	21	1,96%

and biobuthanol in primis), could be filed under all the three key IPC classes C07C “Carbocyclic compounds”, B01J “chemical processes, e.g. catalysis” and C12P “fermentation or enzyme-using processes”.

Technologically, patents assigned to B01J, C07C, and C12P in our patent data embody the core technologies for the emerging lignocellulosic biofuels production technologies and methods, and are likely to continue to have high growth rates in the near future. C01L class includes a significant number of patents relevant to lignocellulosic biofuels, but it remains a challenge to differentiate biofuels patents from fuel additives and lubricants in the class.

Table 4 describes an emerging technology that clearly concentrates on a few key technologies, of which the most important ones relate to exploiting the advantages of lignocellulosic biofuel as a high-octane fuel suitable for mixed-fuel combustion engines. Investment in research and development of lignocellulosic biofuel technologies is likely to continue at the current (or an increased rate) in the near future, and it is likely that the key technologies will demonstrate more pronounced growth dynamics.

4.4. Technological nature of lignocellulosic biofuels patents

An analysis on how the different key lignocellulosic biofuel technologies are related casts light on the nature of the emerging technological solutions, as well as on how different technical solutions relate to the overall production technologies. This type of analysis, in particular, shows to what degree solutions focus on specific technological areas.

By using traditional analysis of co-classification of patent technology classes, Table 5 shows what percentage of patents classified in the leading IPC classes have been co-classified to other key technology classes. Of the six largest technology classes detailed in the rows, C07C, C10L, and C08F have received only moderate co-classifications, whereas B01J, C12P and C10G have received relatively modestly co-classifications.

Of all classes, only C07C and B01J establish a clear pair in a way that reveals that much of B01J classified technology is dependent on C10G technologies. Given that these two technology classifications also demonstrate the highest growth in annual patenting in recent years, it is likely that many of the new critical solutions in lignocellulosic biofuels are being developed in this nexus. However, it is important to note that a significant portion of patents co-classified in C07C and B01J are obtained by the leading incumbents, the Celanese Corporation, Shell, and UOP, as well as, to lesser extent, by focussed new biofuel firms, such as Range Fuels. Furthermore, the majority of these co-classified patents are filed by U.S. based inventors.

Table 5
Technological interrelatedness of patents, 2002–2015.
Source: USPTO, Teqmine Analytics Oy, Authors.

IPC Class	C07C	C10L	B01J	C12P	C08F	C10G	Patents
C07C		9%	19%	5%	3%	4%	431
C10L	13%		6%	4%	0%	9%	304
B01J	66%	14%		6%	4%	10%	127
C12P	17%	10%	6%		6%	4%	127
C08F	12%	1%	5%	7%		0%	105
C10G	19%	27%	13%	5%	0%		101
C07D	80%	15%	15%	11%	4%	10%	80
C13K	17%	5%	6%	23%	2%	0%	65
C08B	23%	2%	8%	13%	6%	2%	53
C02F	6%	0%	14%	0%	0%	2%	51
C08G	68%	9%	6%	15%	35%	0%	34
C08L	21%	10%	7%	7%	62%	0%	29
C10M	39%	75%	0%	0%	21%	4%	28
C07H	60%	4%	28%	24%	12%	4%	25
B01D	41%	14%	14%	0%	9%	18%	22

Note: The table follows whole count: Patent is classified in a class only once. A patent may receive multiple IPC classifications.

While a detailed analysis of the exact nature of the contemporary technological inter-relatedness in lignocellulosic biofuel patents is beyond the scope of this paper, as it would necessitate close examination of the context of up to hundreds of patents, some broad features can be extracted with our patent analysis. Much of the inventive effort appears to focus on three key aspects of lignocellulosic biofuels: exploitation of the high octane content of the raw material, ecological pre-treatment processes for fermentation technologies, and improving the overall fuel yields. Combined, these are indeed inter-related, as they bear upon the technical and engineering realities to arrive at reasonable and economically sufficient cost-benefit ratio in large-scale lignocellulosic biofuel production.

5. Discussion and conclusions

We have demonstrated with Big Data patent analysis how the technological and business environment of lignocellulosic biofuels is in transition, and have argued that such technological intelligence is a valuable decision-support for the strategic management of firms, research and academia, as well as government.

The field of lignocellulosic biofuel technologies is a nascent, emerging technology, characterized by an intense patent race. The overall patenting in the technology has increased about eightfold since 2002, but in selected sub-technology areas, the growth exceeds this. In the near future, the combined annual patent application rate for the technology is likely to be between 150 and 200, with an increasing focus on the biomolecules that may be chemically substituted for major petroleum-based molecules, such as the hydrocarbons benzene and cyclopropane, which are typically filed in patents under the IPC classification C07C.

The intensity of patenting growth has broad implications for the lignocellulosic biofuel community. Most importantly, the IPR space is becoming increasingly densely populated, and individual patents increasingly interdependent, which affects the possibility of obtaining broad exclusive IPR rights in the future, as well as the possibly limiting the scope of past applications or existing granted patents. This development, combined with the fact that certain key IPR holders, such as Celanese or other major corporations, are building large, consolidated IPR portfolios, is likely to have a bearing on assignees' possibilities of exploiting individual or small IPR portfolios. It is likely that full-scale production systems must draw on a broad range of IPR, which are not generally found exclusively in only one portfolio. Thus, we suspect that the industry is likely to increasingly explore collaborative arrangements in order to exploit and develop and IPR in the future.

The globally leading countries in lignocellulosic biofuels are the U.S., Germany, Japan, France and the U.K. These countries are demonstrating substantial advances and investments in research and development, and have typically few or several specialized regional clusters and firm groups active in lignocellulosic biofuels research. A follow-up group consists of Canada, the Netherlands, Belgium and Israel, which all show an increasing effort in the area. The traditional forest industry and technology countries, Sweden and Finland, show relatively modest activity within the technology area.

Because our patent data has been selected to represent the latest available data, i.e. mostly first-time publications of patent applications, we are not able to offer a credible analysis of assignees. However, it is obvious that commercial interest in lignocellulosic biofuel innovation is coming from different groups. The incumbent chemical industry and oil industry firms, such as Celanese, BASF, Shell and Honeywell UOP, are the patenting leaders in the field. A number of smaller new entrants, such as Virent, KIOR or Range Fuels, also play an important role. Research institutes and universities show declining importance, probably because the commercial firms are engaging in an intensifying patent race. Finally, one should note that the incumbent forest industry firms are not central IPR holders, although some of them have single important patents.

As we have documented increased patenting and evident patent races in lignocellulosic biofuels, it is important that future research explores in more detail how IPR rights impact the incentives for innovation, as well as how diffusion of new technologies is affected both in rich and poor countries. Given that we have focussed on the operational aspects of lignocellulosic biofuel innovation and patents, one important avenue for future research is to analyse how the changing nature of innovation and patents affects the business strategy of new entrants, such as research-based start-ups, and industry incumbents. It would also be important to understand in more detail why chemical industries and firms outperform traditional forest products firms in lignocellulosic biofuel invention. Finally, the emergence of technological and production systems of renewable biofuels is a phenomenon linked to broader societal, economic and environmental changes, such as increased demand for energy in developing countries, and future research should also explore the broader relationships between the changing nature of energy demand, biofuels innovation, and political economy.

Acknowledgements

The authors would like to thank Arho Suominen for advice and assistance with the LDA. This research has been supported with research grant (40204/174) from Tekes – The Finnish Funding Agency for Innovation – and with research grant 2014-06082 from Vinnova – the Swedish Innovation system Agency – for the project "RAID: Radical and Incremental Innovation"for the project "RAID: Radical and Incremental Innovation. The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

[1] Bettencourt LMA, Trancik JE, Kaur J. Determinants of the pace of global innovation in energy technologies. PLoS One 2013;8:e67864. http://dx.doi.org/10.1371/journal.pone.0067864.

[2] Klapper R, Upham P, Allison G Who owns the future? Reflections on patenting, private value accrual and societal disbenefit in the context of biofuel technology transfer. Bentham Science Publishers; Available: ⟨http://www.ingentaconnect.com/content/ben/tte/2015/00000002/00000002/art00003⟩.

[3] Lee S, Yoon B, Park Y. An approach to discovering new technology opportunities: Keyword-based patent map approach. Technovation 2009;29:481–97. http://dx.doi.org/10.1016/j.technovation.2008.10.006.

[4] Holladay JE, White JF, Bozell JJ, Johnson D. Top value-added chemicals from biomass -volume II? Results of screening for potential candidates from biorefinery lignin [Internet]. Richland, WA; 2007 Oct. ⟨http://dx.doi.org/10.2172/921839⟩.

[5] Bozell JJ, Petersen GR. Technology development for the production of biobased products from biorefinery carbohydrates—the US department of energy's "Top 10" revisited. Green Chem R Soc Chem 2010;12:539. http://dx.doi.org/10.1039/b922014c.

[6] Valipour M. Future of agricultural water management in Africa. Arch Agron Soil Sci Taylor Francis 2015;61:907–27. http://dx.doi.org/10.1080/03650340.2014.961433.

[7] Kamm B, Gruber PR, Kamm M. Biorefineries - Industrial Processes and Products: Status Quo and Future Directions. Kamm B, Gruber PR, Kamm M, editors. Mörlenbach: Wiley VCH; 2010.

[8] Toivanen H. Learning and corporate strategy: the dynamic evolution of the North American pulp and paper industry, 1860–1960 [Internet] (Ph.D. Diss.). Atlanta: Georgia Instittue of Technology; 2004, [Available] ⟨http://hdl.handle.net/1853/4972⟩.

[9] Toivanen H. Waves of technological innovation: the evolution of the US pulp and paper industry, 1860–2000. Evol Glob Pap Ind 1800–2050 2012; 2012. p. 55–88, [Available] ⟨http://link.springer.com/chapter/10.1007/978-94-007-5431-7_3⟩.

[10] Srirangan K, Akawi L, Moo-Young M, Chou CP. Towards sustainable production of clean energy carriers from biomass resources. Appl Energy 2012;100:172–86. http://dx.doi.org/10.1016/j.apenergy.2012.05.012.

[11] Andresen AD. Innovation systems and natural resources – the case of sugarcane in Brazil. Aalborg, Denmark: Aalborg University; 2011.

[12] Olah GA, Goeppert A, Prakash GKS. Beyond oil and gas: the methanol economy [Internet]; 2011. Available: ⟨https://books.google.com/books?hl=sv&lr=&id=WxwktXYR-W4C&pgis=1⟩.

[13] Chisti Y. Biodiesel from microalgae. Biotechnol Adv 2007;25:294–306. http://dx.doi.org/10.1016/j.biotechadv.2007.02.001.

[14] De Wild PJ, Huijgen WJJ, Gosselink RJA. Lignin pyrolysis for profitable lignocellulosic biorefineries. Biofuels, Bioprod Bioref 2014;8:645–57. http://dx.doi.org/10.1002/bbb.1474.

[15] Novotny M, Laestadius S. Beyond papermaking: technology and market shifts for wood-based biomass industries – management implications for large-scale industries. Technol Anal Strateg Manag Routledge 2014:1–17. http://dx.doi.org/10.1080/09537325.2014.912789.

[16] Leu H-J, Wu C-C, Lin C-Y. Technology exploration and forecasting of biofuels and biohydrogen energy from patent analysis. Int J Hydrog Energy 2012;37:15719–25. http://dx.doi.org/10.1016/j.ijhydene.2012.04.143.

[17] Hsu C-W, Chang P-L, Hsiung C-M, Wu C-C. Charting the evolution of biohydrogen production technology through a patent analysis. Biomass- Bioenergy 2015;76:1–10. http://dx.doi.org/10.1016/j.biombioe.2015.02.035.

[18] Albino V, Ardito L, Dangelico RM, Messeni Petruzzelli A. Understanding the development trends of low-carbon energy technologies: a patent analysis. Appl Energy 2014;135:836–54. http://dx.doi.org/10.1016/j.apenergy.2014.08.012.

[19] Klepper S. Industry life cycles. Ind corp chang 1997;6:145–82, Available: ⟨http://icc.oxfordjournals.org/content/6/1/145.abstract⟩. Accessed 15 March 2016.

[20] Cohen WM, Klepper S. A reprise of size and R & D. Econ J 1996;106:925–51, [Royal Economic Society, Wiley] ⟨http://www.jstor.org/stable/2235365⟩.

[21] Hall BH, Jaffe A, Trajtenberg M. Market value and patent citations on. RAND J Econ 2005;36:16–38. http://dx.doi.org/10.2307/1593752.

[22] Newman NC, Porter AL, Newman D, Trumbach CC, Bolan SD. Comparing methods to extract technical content for technological intelligence. J Eng Technol Manag 2014;32:97–109. http://dx.doi.org/10.1016/j.jengtecman.2013.09.001.
[23] Suominen A. Analysis of technological progression by quantitative measures: a comparison of two technologies. Technol Anal Strateg Manag Routledge 2013;25:687–706. http://dx.doi.org/10.1080/09537325.2013.802930.
[24] Venugopalan S, Rai V. Topic based classification and pattern identification in patents. Technol Forecast Soc Change 2014. http://dx.doi.org/10.1016/j.techfore.2014.10.006.
[25] Hoffman MD, Blei DM, Wang C, Paisley J. Stochastic variational inference. J Mach Learn Res. JMLR.org 2013;14:1303, [1303-1347–1347. Available] ⟨http://dl.acm.org/citation.cfm?id=2567709.2502622⟩.