Patent data for comparative study: Case study of top Aspirants in Bioinformatics Industry

Mago, Neha; Deshpande, Nishad

Patent data for comparative study: Case study of top Aspirants in Bioinformatics Industry

Neha Mago neha.mago19@gmail.com
CSIR-Unit for Research and Development of Information Products (URDIP), India
Nishad Deshpande nishad@urdip.res.in
CSIR-Unit for Research and Development of Information Products (URDIP), India

Abstract: Innovation and technology are considered as a subject of success and achievement to the firm. The comparative study represents an essential procedure to identify the innovation and technological capabilities of major players involved in the bioinformatics related inventions. The aim of the research is to map out the top firms and identify their strategically important technologies. In view of this, the comparative analysis of major firms in bioinformatics industry is carried out using patent information. Herein the top three assignees are considered and based on this further analysis is performed. The top companies' trend suggests that Thermo Fisher Scientific Inc. is the major player in bioinformatics research. Thus, we have tried to develop an overview on their patenting trend and important concerned areas of research. Also, our results indicate that the application of computational tools is being utilized for most of the research areas like the study of genomics and proteomics, sequence categorization and their structural prediction.

Keywords: Patent, Innovation, Comparative Advantage, Company, Bioinformatic, Industry, Computational.

Resumen: La innovación y la tecnología se consideran un tema de éxito y logro para la empresa. El estudio comparativo representa un procedimiento esencial para identificar la innovación y las capacidades tecnológicas de los principales actores involucrados en las invenciones relacionadas con la bioinformática. El objetivo de la investigación es trazar un mapa de las principales firmas e identificar sus tecnologías estratégicamente importantes. En vista de esto, el análisis comparativo de las principales empresas en la industria de la bioinformática se lleva a cabo utilizando información de patentes. Aquí se consideran los tres principales cesionarios y se basa en este análisis adicional. La tendencia de las principales compañías sugiere que Thermo Fisher Scientific Inc. es el principal actor en la investigación bioinformática. Por lo tanto, hemos tratado de desarrollar una visión general de su tendencia de patentamiento y áreas de investigación importantes. Además, nuestros resultados indican que la aplicación de herramientas computacionales se está utilizando para la mayoría de las áreas de investigación como el estudio de la genómica y la proteómica, la categorización de secuencias y su predicción estructural.

Palabras clave: Patente, Innovación, Ventaja Comparativa, Empresa, Bioinformática, Industria, Computacional.

INTRODUCTION

Various methods of technological expansion in bioinformatics research over last 15 decades has given rise to the interest in protecting such creative ideas to uplift the research area. The industries which are dominant on the bioinformatic sector has delivered successful inventions for predicting functions of unknown genes and protein, sequence related analysis, and mathematical or statistical based simulations (Perez-Iratxeta et al., 2007). They have also, provided solutions to many diseases to the world, wherein scientist and researchers work-in-parallel with the vision of the industry (Chow et al., 2001). The bioinformatic industry has supported the pharmaceutical and biotechnological industries to contribute to the analysis and management of data obtained from genomics and proteomics. Such analytical tools, software and hardware products are vitally important, especially in the process of drug discovery and development. The products are highly effective and accurately analyse the data generated from high-throughput technologies such as DNA microarrays technologies (Chang et al., 2004). Thus, the innovation and technology are equally important for the growth of an industry. The advancement in the research and development activity gives rise to the increase in technological development and innovations, happening in product and process technologies (Santos et al., 2015). However, it is advisable to perform technology and state-of-art analysis before bringing new inventions in productive form. The analysis provides a global overview regarding the development of certain technology and the major companies involved.

INNOVATION AND BIOINFORMATICS

The development observed from the last so many decades in the bioinformatic sector has fascinated industries that mainly belongs to biotechnology and pharmaceutical sector (Chang et al., 2004). The most commonly noticed shift have been from the process of drug discovery to in-silico structure based drug design (Choi et al., 2016). The application of computer technology for the management of biological information delineates bioinformatics. Wherein the activity of biological computing makes various system to work efficiently to easily access, manage and retrieve biological information. It resolves biological problems in the areas of DNA and Protein related analysis which includes sequence categorization, structural and functional prediction, protein interaction studies and various modeling pathways (Ouzounis et al., 2003). Furthermore, the next-generation sequencing generates an enormous amount of data which requires storing and transforming the information into value in new ways (de Brevern et al., 2015).

Value of patent information

Patents are indeed unique and essential information source holding new technology information. They provide an insight into industry developments, technology advancement, and information about new invention in detail. It is considered as a well-categorized technical documents updated till date on new and innovative technologies.

Worldwide, the patent document is filed in accordance with the regional or national patent system. Wherein the information is indexed in standardized form uncovering the background and details of the invention along with its limitation and application. A patent application takes almost 18 months to get published prior to the date of filing and an applicant may use its protection right for at most 20 years considered from the date of filing. So, there is a time lag in patent procedure between the filing and publication of the patent application. The patent referred as the technolegal document, which assists in defining the knife-edge of technological information covering innovative product or process, and consequentially it reflects the success or failure of a company. For this reason, the protection of any sort of creative thinking is vital in the form of patents. The time spent on strategy at any juncture of a patent intelligence is important for the progress and success of revealed technical information (Hong et al., 2007). However, the patenting trend and information covered within the technical documents help in the identification of new and emerging competitors' along with their technology development. With the patent datasets and information, the competitive intelligence of major aspirants can be observed (Katila et al., 2000). The comparative analysis of various companies denotes the strategically important technologies utilised respectively and also, identifies if they are with the pace of technology (Afuah, 2002). So, this paper covers the comparative analysis of the firms' to develop insights on their patenting trends, main focus areas of research and shedding a light on their strategy to compete with the day-to-day technology advancement.

METHOD

In order to achieve the above objective, it was required to extract the patent data for which an effective search was carried out. The executed search strategy was based on technology-based keywords, IPC (International Patent Classification) and CPC (Cooperative Patent Classification) system in relation to bioinformatics related invention. The most relevant patent applications were retrieved were until December 2015. As the innovation and technology are equally important for the growth of an industry [4]. Thus, in this paper, we will examine the analogy of top leading players in the bioinformatics industry. First, we will observe the patenting trend and major aspirants of bioinformatics related inventions. Later, we will perform the comparison with the major active players and explore their advancement and development under bioinformatics sector by executing correlation between them.

RESULTS AND DISCUSSION

The comparative study for top leading companies was performed to provide strategically important technologies in bioinformatics industry.

The analysis will provide comparison between the firms by the way of gauging their innovation and technological capabilities.

Patenting trend

Figure 1 depicts the distribution of patent applications filed annually where the patent filing year was considered for the analysis. The earliest patenting activity was seen in the year 1990s. The steady growth and increase in patent trend was observed after the year 2002. The pickup of patenting activity was recorded in the year 2013 and onwards, where the number of patent applications were maximum. During this period, the peak indicated the increase in research interest in bioinformatics industry. Also, the number of patent applications during the year 2014 and 2015 is incompletely represented due to the procedural time lag in patent filing as patent applications are published only after a period of 18 months from the date of filing. Also, it is worth noting that over the period of last 15 years there is a continued research and innovative growth in the field.
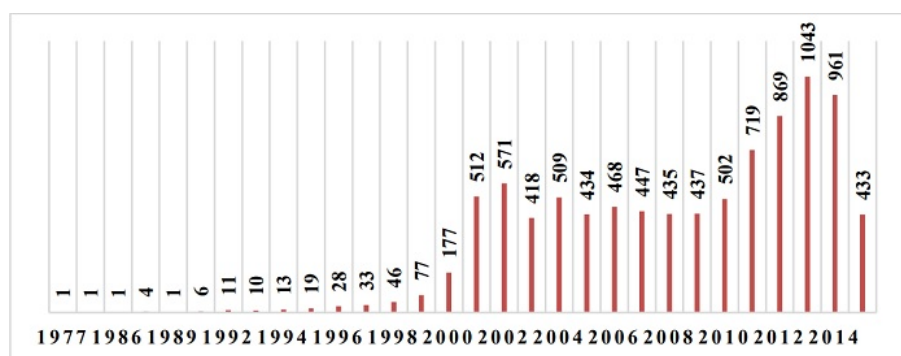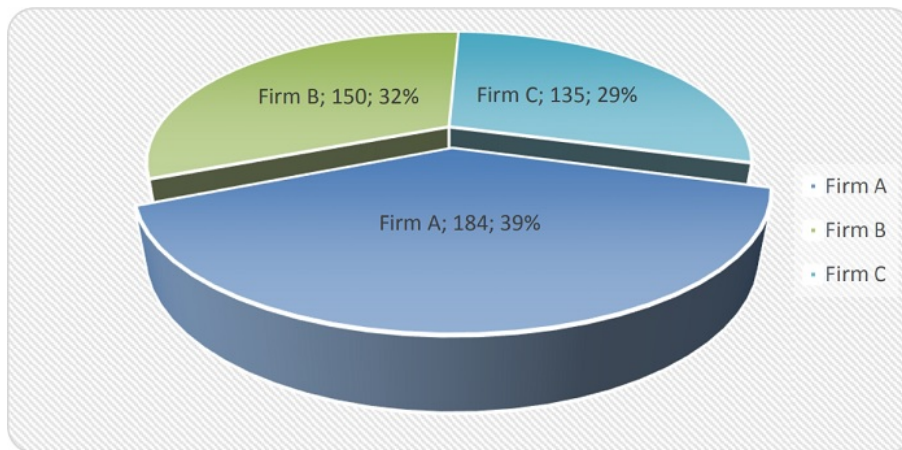

Figure 1
Overall Patent Filing Trend.

Comparative analysis for top assignees

The R&D emphasis by competitors and their strategies in the specific technological field can be interpreted from the companies patenting activity. The companies' patenting activity and competitive position are analyzed from the retrieved patent datasets. The top three applicants are considered in respect of the number of patents filed every year. Their patent filing activity and technological advancement is observed. The parameters used to compare the top applicants/assignees development, are discussed as below;

Patent share

The number of patents held by assignee indicates the dominant players in bioinformaticsrelated research and their extent to exploit a certain technology. The top three assignees represented as Thermo Fisher Scientific Inc. (Firm A), Hitachi Ltd. (Firm B) And IBM Corp. (Firm

C) were considered for further analysis. Figure 2 reveals the competitive position within the bioinformatics industry which denote the share of patents held by the top three firms. Firm A, has filed the maximum number of patent applications. The patent filing activity by Firm A is the highest after the year of 2012. The company works maximum in genomics, proteomics, and sequence related research to provide numerous health-related solutions and has become a premier leader in life science sector. The success behind the growth of a company lies in major merger and acquisition with the world's best brand like Applied Biosystems, Life technologies Corp, Invitrogen etc.

Figure 2
Comparison between top applicants based on the count of patent filed.

The Table 1 represents the count distributed among the topmost technologies covered in the field of bioinformatics and related inventions. Firm A covers the majority of share in research area which relates to sequence analysis, determining or predicting the function of genes and proteins, and utilization of computational technologies. Whereas, the patent application count of Firm B represents that the sequence related analysis and informatics are the majorly focused areas. Similarly, Firm C works mainly in the same research area as Firm A and Firm B but there patent filing was activity observed to be significantly less as compared to other two. The complete count distribution of patent filing application in certain technology indicates that Firm A is a clear leader in the bioinformatics industry.

Table 1

List of number of filed patent categorised in respect of majorly used IPC class.

Majorly Used IPC class	Description	Firm A	Firm B	Firm C
G06F 19/12	Modelling or simulations in systems biology	11	6	17
G06F 19/14	Phylogeny or evolution	3	5	6
G06F 19/16	Molecular structure	3	12	21
G06F 19/18	Functional genomics or proteomics	59	29	24
G06F 19/20	Hybridization or gene expression	37	36	11
G06F 19/22	Sequence comparison	71	51	63
G06F 19/24	Machine learning, data mining or biostatistics	41	22	41
G06F 19/26	Data visualization	24	31	6
G06F 19/28	Programming tools or database system	36	44	25

Bi-variate correlation

The Bivariate Pearson Correlation twotailed significance test is performed to compare the patent filing activity among top three companies. The Pearson correlation is capable of detecting linear associations between variables and it was carried out using PSPP software. The data represented in table 1 is used for the calculation were majorly utilized technologies were distributed. The correlation comes in the range of -1 to +1 where the range indicates the strength of the relationship. Figure 3 gives all the correlation and two-tailed significance value among the top three firms. It represents the lowest index of correlation for Firm B and Firm C. This brings out important facts in respect to specific technologies. The correlation between the three firms showed that firm C is significantly less related and signifies that it is far apart in the research area of bioinformatics industry. On the other hand, firm A and Firm B are statistically closely related which represents that the two firms are sharing interests and focus in the similar research area.

```
CORRELATIONS

CORRELATION
    /VARIABLES =  FirmA FirmB FirmC
    /PRINT = TWOTAIL SIG.
```

Correlations

		FirmA	*FirmB*	*FirmC*
FirmA	*Pearson Correlation*	1.00	.81	.72
	Sig. (2-tailed)		.008	.029
	N	9	9	9
FirmB	*Pearson Correlation*	.81	1.00	.53
	Sig. (2-tailed)	.008		.144
	N	9	9	9
FirmC	*Pearson Correlation*	.72	.53	1.00
	Sig. (2-tailed)	.029	.144	
	N	9	9	9

Figure 3
The bivariate correlation calculated to identify the
correlation between top 3 applicants using PSPP Software

CONCLUSION

The analysis of patent information extracted for the bioinformatic related inventions is shedding a light on the topmost innovator firms. It reveals the patent filing trend in the bioinformatic sector. The increase in trend is observed in recent years 2013 and onwards. However, the Thermo Fisher Scientific Inc. represented as Firm A in the paper, is recorded as a top leading patent assignee holding a maximum number of patent applications in the corresponding year. Further, the study has shown the comparison between the top leading players and considered two main parameters for the analysis. One is observing the patent held by each of them and other is finding the statistical relationship based on the various technological areas covered under bioinformatics industry. In view of mentioned parameters, it has been determined that Thermo Fisher Scientific Inc. and Hitachi Ltd. are working in the similar research areas. It depicts that both the companies followed a similar patenting strategy to focus on the majorly concerned research technologies like functional genomics and proteomics, sequence related research, and protein structure predictions.

The assignee analysis suggests the best strategically important technologies to the industry. Thus, concluding that Thermo Fisher Scientific Inc. is an innovator firm overriding in the industry providing new and innovative technologies which cover processes and products. In last five years, they have been merging and acquiring the top brands like Applied Biosystems, Life technologies Corp, Invitrogen etc. So, in technolgy-driven age the innovation and creativity holds the success of firm where patenting strategy is essential. Alongwith, alignment of

such IP strategy with the company's objectives and business plan is also important. Wherein relevant decisions are made for technology exploitation in best possible way which leads to the development and success of a company.

REFERENCES

Afuah, A. (2002). Mapping technological capabilities into product markets and competitive advantage: the case of cholesterol drugs. Strategic Management Journal, 23(2), 171-179.

Chang, C. (2004). The Current Bioinformatics Analytical Software Landscape. DDW Drug Discovery World, 5(3), 30-36.

Choi, S., Macalino, S. J., Cui, M., & Basith, S. (2016). Expediting the Design, Discovery, and Development of Anticancer Drugs using Computational Approaches. Current medicinal chemistry.

Chow, M., & Fernandez, D. (2001, October). Intellectual property strategy in bioinformatics. In Proceedings of the Virtual Conference on Genomics and Bioinformatics, North Dakota State University (pp. 15-16).

de Brevern, A. G., Meyniel, J. P., Fairhead, C., Neuvéglise, C., & Malpertuy, A. (2015). Trends in it innovation to build a next generation bioinformatics solution to manage and analyse biological big data produced by NGS technologies. BioMed research international, 2015.

Hong, S. (2007). The magic of patent information. World Intellectual Property Organization, www. wipo. int/sme/en/documents/patent_information. htm.

Katila, R. (2000). Using patent data to measure innovation performance. International Journal of Business Performance Management, 2(1-3), 180-193.

Ouzounis, C. A., & Valencia, A. (2003). Early bioinformatics: the birth of a discipline—a personal view. Bioinformatics, 19(17), 21762190.

Perez-Iratxeta, C., Andrade-Navarro, M. A., & Wren, J. D. (2007). Evolving research trends in bioinformatics. Briefings in Bioinformatics, 8(2), 88-95.

Santos, J., Zawislak, P., Franzoni, G., & Vieira, H.(2015). Searching for a Path: A Bibliometric study on Innovation and Technological Capabilities. International Journal of Innovation, 3(2), 54-66.