



Universidade Estadual Paulista “Júlio de Mesquita Filho” - Unesp
Faculdade de Filosofia e Ciências – Campus de Marília
Programa de Pós-Graduação em Ciência da Informação

FERNANDA BOCHI

**Caracterização de um domínio tecnológico pelas análises relacionais de
citação: um estudo nas patentes sobre células-tronco**

MARÍLIA

2023

FERNANDA BOCHI

Caracterização de um domínio tecnológico pelas análises relacionais de citação: um estudo nas patentes sobre células-tronco

Tese apresentada ao Programa de Pós-Graduação em Ciência da Informação da Universidade Estadual Paulista, como requisito parcial para a obtenção do título de Doutor em Ciência da Informação.

Área de Concentração: Informação, tecnologia e conhecimento.

Linha de pesquisa: Produção e Organização da Informação

Orientadora: Prof^a Dr^a Maria Cláudia Cabrine Grácio.

MARÍLIA

2023

B664c Bochi, Fernanda
Caracterização de um domínio tecnológico pelas análises relacionais de citação : um estudo nas patentes sobre células-tronco / Fernanda Bochi. -- Marília, 2023
140 f. : il., tabs.

Tese (doutorado) - Universidade Estadual Paulista (Unesp), Faculdade de Filosofia e Ciências, Marília
Orientadora: Maria Cláudia Cabrini Grácio

1. Patentes. 2. Bibliometria. 3. Acoplamento bibliográfico. 4. Cocitação. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp.
Biblioteca da Faculdade de Filosofia e Ciências, Marília. Dados
fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

Universidade Estadual Paulista - UNESP
Faculdade de Filosofia e Ciências
Programa de Pós-Graduação em Ciência da Informação

FERNANDA BOCHI

**Caracterização de um domínio tecnológico pelas análises relacionais de
citação: um estudo nas patentes sobre células-tronco**

Tese apresentada ao Programa de Pós-Graduação em Ciência da Informação da
Universidade Estadual Paulista, como requisito parcial para a obtenção do título de
Doutor em Ciência da Informação.

Banca Examinadora

Titular 1. Orientadora: Prof^a. Dr^a. Maria Cláudia Cabrini Grácio (UNESP, Marília).

Titular 2. Prof. Dr. Daniel Martínez-Ávila (UNESP, Marília).

Titular 4. Prof^a. Dr^a. Ely Francina Tannuri de Oliveira (UNESP, Marília).

Titular 3. Prof. Dr. Leandro Innocentini Lopes de Faria (UFSCar, São Carlos).

Titular 5. Prof^a. Dr^a. Ana Maria Mielniczuk de Moura (UFRGS, Porto Alegre).

Marília, 13 de abril de 2023

AGRADECIMENTO

Dedicar alguns minutos para trazer à memória todos aqueles que foram importantes para mim durante o doutorado é muito satisfatório. Este registro é feito com o coração cheio de alegria numa singela forma de demonstrar gratidão e reconhecer a importância que cada pessoa tem na minha vida.

Sou grata, primeiramente a Deus, porque dEle, por Ele e para Ele são todas as minhas coisas.

Sou grata à minha mãe, por ser força e incentivo nos momentos de angústia. Grata aos meus irmãos, em especial ao Fabiano, por me ajudar e encorajar. Às minhas irmãs Lígia, Cristina e Cláudia pelas orações, palavras de ânimo e mimos inesperados quando me via solitária nas tardes de domingo em Marília. Minha eterna gratidão à minha cunhada Patrícia, por dedicar muitas horas do seu tempo para me ajudar e por acreditar que eu poderia ir além. Gratidão pelos meus sobrinhos Felipe e Beatriz, por me doarem alegria e amor. A Juliana Furlan, amiga biomédica que a vida me deu e que sempre me ajuda a compreender termos tão técnicos.

Sou grata à minha incansável orientadora Prof^a Dr^a Maria Cláudia Cabrini Grácio que sempre, tão humana e dedicada, me desafiou a acreditar em mim; Viu habilidades que não conseguia ver. Aconselhou-me e encorajou-me a seguir em frente. Fazendo tudo com uma naturalidade e leveza admirável. Meu respeito e admiração crescem todas as vezes que temos alguns minutos de diálogo, mesmo quando acontecem apenas pelo WhatsApp. Sempre digo e agora registro, que espero que possamos seguir caminhando juntas.

Ao querido Prof. Dr. Rene Faustino Gabriel Jr., minha eterna gratidão pelo suporte com os dados de pesquisa, pelas palavras de incentivo e por me impulsionar a seguir pesquisando. Gratidão por gerar mais dúvidas aos meus questionamentos, mostrando que podemos pensar além. Considero um grande provocador de questões e um grande amigo que a vida me deu.

Sou grata à Prof^a Dr^a Ana Maria M. de Moura, por acreditar que eu não deveria ficar só no mestrado. Por me fazer pensar longe e por me incentivar. Acredito que nossa relação já não ocupe apenas o espaço acadêmico, pois foi edificada sobre bases sólidas de amizade e confiança. Você é um ser humano incrível. Obrigada por tanto!

Aos membros da banca avaliadora sou grata pelas sugestões e discussões que iniciaram na qualificação e seguiram na defesa, enriquecendo o trabalho: Prof. Dr. Daniel Martínez-Ávila, Prof^a Dr^a Ely Francina Tannuri de Oliveira, Prof. Dr. Leandro Innocentini Lopes de Faria, Prof^a Dr^a Ana Maria Mielniczuk de Moura.

Aos colegas de pós-graduação, em especial a Deise Deolindo pela grande e inesquecível ajuda no começo de tudo isso. Ao meu colega e amigo Rafael Castanha, pelo apoio com os dados e por ser incansável em ajudar. Espero que a vida brilhe sempre para você! Às minhas amadas e doces colegas/amigas Bianca S. de Mira e Natalia Delbianco, por serem tão divertidas e dedicadas. Vocês tornam o processo mais leve. À amiga e colega Carla Hilário, pela destreza nas discussões do grupo, tornando tudo mais rico. Não posso deixar de mencionar que você e o Thales são excelentes anfitriões e que a minha rápida passagem pela UEL foi muito agradável porque vocês estavam lá.

Ao querido amigo e colega Jean Fernandes Brito, pelas longas e incansáveis conversas sobre tudo nas idas ao mercado. Pelos sucos que tornavam o dia quente mais refrescante e mais leve. Você foi um grande presente que a vida me deu. À minha amiga e colega Regina, que mesmo longe sempre se fez presente. Incansável em ser gentil. Obrigada! Aos colegas Clayton, Larissa, Mirelys e Taciana, pelas conversas nas horas vagas e trocas nas aulas de metodologia. Aos amigos e colegas de trabalho dos grupos de pesquisa Estudos Métricos em Informação (EMI-Unesp) e Núcleo de Estudos em Ciência, Inovação e Tecnologia (NECIT-UFRGS).

Agradeço as amigas Dariele e Mônica, por orarem por mim e por terem, sempre, uma palavra de ânimo e benção.

À Universidade Estadual Paulista, pelos excelentes professores e pela infraestrutura oferecida, permitindo que eu realizasse o sonho de uma vida.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

“...pense na ciência como uma dádiva do Criador. A curiosidade que nos foi inculcada para entender como o universo funciona pode inspirar ainda mais admiração ao Criador. De forma alguma, esta dádiva poderia ser uma ameaça para Deus, o autor de tudo. Celebre o que a ciência pode nos ensinar. Pense na ciência como uma forma de adoração.”

Francis Collins

BOCHI, F. **Caracterização de um domínio tecnológico pelas análises relacionais de citação**: um estudo nas patentes sobre células-tronco. 2023. Tese (Doutorado em Ciência da Informação) – Faculdade de Filosofia e Ciência, Universidade Estadual Paulista, Marília, 2023.

RESUMO

A análise de citação uni-variada e relacional, bem como a análise de domínio, são técnicas que demonstram relevância para compreensão das frentes tecnológicas de diversos campos. Assim, a fim de caracterizar a proximidade das identidades tecnológicas das patentes e suas imagens diante da comunidade tecnológica em células-tronco, a pesquisa se desenvolveu tendo como problemática: Qual a contribuição da análise conjunta de métodos bibliométricos uni-variados e relacionais de citação para caracterização de um domínio tecnológico formado por patentes envolvendo células-tronco? Assim, conjectura-se que os métodos bibliométricos uni-variados e relacionais de citação são capazes de caracterizar e identificar domínios tecnológicos no âmbito das patentes. A pesquisa objetiva analisar a contribuição dos métodos de análise uni-variadas e relacionais de citação, aplicados às patentes, para a identificação e caracterização de domínios científico-tecnológicos, especificamente: identificar o referencial tecnológico (citações as patentes) presente no universo das patentes envolvendo células-tronco, analisar o referencial científico (citações à literatura não-patente) adotado no conjunto de patentes de células-tronco concedidas, identificar os agrupamentos da produção tecnológica relacionadas a células-tronco, caracterizados pelas proximidades tecnológicas e científicas das patentes indexadas na base DII, evidenciados pelas similaridades nas suas referências e mensuradas pelo acoplamento bibliográfico, descrever os agrupamentos tecnológicos e científicos identificados a partir das imagens (citações simultâneas) das patentes em células-tronco e reconhecer os métodos bibliométricos de acoplamento bibliográfico e cocitação entre patentes para identificação de domínios científicos-tecnológicos. Para atender os objetivos estabelecidos empregou-se a patentometria, que tem como principal atributo a análise do uso e das características dos documentos de patentes, associada aos procedimentos consolidados da bibliometria, mais especialmente estudos de cocitação e acoplamento bibliográfico, tendo como corpus 144 patentes indexadas na base de dado Derwent Innovation Index, entre o período de 2016 a 2020. Em relação à análise do referencial científico, observou-se que há um número significativo de autores pouco citados e poucos autores muito citados. No entanto, a pesquisa mostrou que mesmo o referente mais citado (74 citações) não apresentou uma representatividade quantitativa, correspondendo a apenas 0,51%. O estudo indicou, quanto à análise de acoplamento, que as patentes com maior proximidade de acoplamento, com links de acoplamento entre 39 (GB2531454) e 73 (WO2017023801), apresentam uma similaridade tecnológica e temática na área de Biotecnologia aplicada à Microbiologia e Química, uma vez que todas apresentam no seu rol de classes a classificação internacional de patente C12N, relativa à microrganismos ou enzimas, suas composições, mutação ou engenharia genética e meios de cultura. O acoplamento bibliográfico de inventores utilizou, como unidade de análise, as referências à literatura científica. Na análise, foi possível observar diversos domínios representados por temas e grupos de indivíduos. Entres os resultados, identificou-se um inventor/pesquisador muito presente: Feng Zhang, com patentes relevantes no tema edição genômica e CRISPR e colaboração científica e tecnológica

com coautores que indicam relacionamento de genealogia acadêmica. Na análise de cocitação, a partir dos referentes, foi possível identificar 5 cluster bem definidos com domínios temáticos variados que se relacionam com os demais clusters. Neste resultado, mais uma vez foi possível identificar a presença de engenharia genética voltada à edição genômica, mais especificamente a técnica CRISPR. Por sua vez, por meio da análise das redes de cocitação, identificaram-se 23 clusters que caracterizam os domínios epistêmicos que congregam as correntes tecnológicas em que atuam os inventores em células-tronco evidenciados na rede de acoplamento das patentes citantes. Os resultados apresentados na pesquisa validaram a hipótese apresentada, no entanto, considerando algumas limitações identificadas no decorrer do estudo, propõe-se a aplicação desta metodologia aliada a estudos de modelagem de tópicos, bem como organização do conhecimento e recuperação da informação.

Palavras-chave: patentes; análise de citação uni-variada; análise de citação relacional; patentometria; análise de acoplamento; análise de cocitação; análise de domínio.

ABSTRACT

Univariate and related citation analyses, as well as domain analysis, are techniques that demonstrate relevance in understanding the technological fronts of different fields. To characterize the proximity of the technological identities of the patents and their images before the technological community in stem cells, the following research has been developed using a certain problematic: What is the contribution of the joint analysis of univariate bibliometric and relational citation methods in the characterization of a technological domain formed by patents involving stem cells? It is conjectured that univariate and relational bibliometric citation methods are capable of both characterizing and identifying technological domains within the scope of patents. This research aims to analyze the contribution of univariate and relational citation analysis methods, applied to patents, for the identification and characterization of scientific-technological domains, in this specific context: to identify the technological reference (citations to patents) present in the field of stem cells, to analyze the scientific framework (citations to non-patent literature) adopted in the set of granted stem cell patents, to identify groups of technological production related to stem cells, characterized by technological and scientific proximity to the patents indexed in the DII database, evidenced by similarities in their references and measured by bibliographic coupling, to describe the technological and scientific groupings identified from the images (simultaneous citations) of patents on stem cells and to recognize the bibliometric methods of bibliographic coupling and co-citation between patents to identify scientific-technological domains. To meet these objectives, patentometrics has been used, whose main attribute is the analysis of the use and characteristics of patent documents associated with consolidated procedures of bibliometrics, more especially studies of co-citation and bibliographic coupling, having as a corpus 144 patents indexed in the Derwent Innovation Index database between 2016 and 2020. Regarding the analysis of the scientific reference, it has been observed that there is a significant number of authors little cited and few authors very much so. However, the research showed that even the most cited referent (74 citations) did not present a quantitative representation, corresponding to only 0.51%. The study indicated, referring to the coupling analysis, that the patents with the greatest proximity of coupling, with coupling links between 39 (GB2531454) and 73 (WO2017023801), present a technological and thematic similarity in the area of Biotechnology applied to Microbiology and Chemistry, since all of them have in their roll of classes the international patent classification C12N, related to microorganisms or enzymes, their compositions, mutation or genetic engineering and culture media. The bibliographic coupling of inventors used references to scientific literature as the unit of analysis. In the analysis, it was possible to observe several domains represented by themes and groups of people. Among the results, an inventor/researcher very present in the results was identified. Feng Zhang has presented relevant patents on the topic of genomic editing and CRISPR and has presented a scientific and technological collaboration with co-authors that indicate an academic genealogy relationship. In the co-citation analysis based on the referents, it was possible to obtain five well-defined clusters that had different thematic domains, but that connected to the other clusters. In this result, it was once again possible to identify the presence of genetic engineering aimed at genomic editing, more specifically the CRISPR technique. In turn, through the analysis of the co-citation networks, it was identified that the twenty-three clusters represent epistemic domains that characterize technological currents in which the inventors of technological domains in stem cells act, evidenced in the coupling network of the citing

patents. The results presented in the research validated the presented hypothesis. However, considering some limitations identified during the study, it is proposed that the application of this methodology is allied to topical modeling studies, as well as to knowledge organization and information retrieval.

Keywords: patents; univariate citation analysis; relational citation analysis; patentometrics; coupling analysis; cocitation analysis; domain analysis.

LISTA DE FIGURAS

Figura 1 – Folha de rosto da patente US2016208243-A1.....	30
Figura 2 – Exemplo de família de patentes.....	32
Figura 3 – Estrutura da Classificação Internacional de Patente.....	34
Figura 4 – Relação de citação e referência entre os documentos A, B, C e D.....	44
Figura 5 – Artigos A e B acoplados bibliograficamente pelos documentos 3, 4, 5 e 6	57
Figura 6 – Ilustração do acoplamento bibliográfico entre patentes por meio do compartilhamento de referências entre elas.....	59
Figura 7 – Documentos 3 e 4 cocitados em função dos artigos citantes A, B e C.....	60
Figura 8 – Ilustração da rede de acoplamento bibliográfico e de cocitação entre patentes	64
Figura 9 – Padronização das famílias de patentes localizadas no campo PN.....	70
Figura 10 – Recorte do processo de transformação dos dados em matriz.....	71
Figura 11 – Identificação dos campos utilizados para leitura dos dados da rede de cocitação.....	72
Figura 12 – Organização das patentes citantes e suas referências citadas em colunas.....	72
Figura 13 – Recorte dos dados gerados no software R para elaboração da rede de acoplamento bibliográfico	73
Figura 14 – Recorte da planilha contendo os nós e as ligações (links) para criar o arquivo .net.....	74
Figura 15 – Exemplo de erros identificados e que foram substituídos pela referência adequada, seguindo a norma da APA.....	75
Figura 16 – Ilustração dos problemas encontrados nos documentos de patentes pela falta de padronização das informações.....	79
Figura 17 – Processo de busca realizado no ORCID.....	80

Figura 18 – Rede de acoplamento bibliográfico entre os inventores que comportam o domínio do estudo.....	96
Figura 19 – Rede de acoplamento tecnológico entre as patentes pertencentes ao domínio analisado.....	100
Figura 20 - Rede de cocitação da literatura científica referenciada nas patentes....	105
Figura 21 – Rede de cocitação das patentes citadas pelas 144 patentes do corpus da pesquisa.....	109

SUMÁRIO

1	INTRODUÇÃO	16
2	PATENTES E INOVAÇÃO: ELEMENTOS HISTÓRICOS E CONCEITUAIS ..	27
2.1	ELEMENTOS QUE COMPÕEM AS PATENTES	30
2.2	LEGISLAÇÕES	35
2.3	TECNOLOGIA EM CÉLULAS-TRONCO.....	37
3	ANÁLISE DE DOMÍNIO	40
4	ANÁLISE DE CITAÇÃO	44
4.1	INDICADORES METRICOS.....	48
4.2	IDENTIDADE E IMAGEM DE CITAÇÃO	53
4.3	ANÁLISES RELACIONAIS DE CITAÇÃO	56
4.3.1	Acoplamento bibliográfico	57
4.3.2	Cocitação	60
5	METODOLOGIA	66
5.1	ABORDAGEM E TIPO DA PESQUISA	66
5.2	DELINEAMENTO DA PESQUISA.....	66
5.3	ESCOLHA DA BASE DE DADOS E ESTRATÉGIA DE BUSCA.....	67
5.3.1	Fonte de Coleta dos Dados	67
5.3.2	Procedimento para Coleta dos Dados	68
5.3.3	Tratamento e Análise dos Dados	70
5.4	LIMITAÇÃO DO ESTUDO	79
5.5	PLANO DE GESTÃO DE DADOS.....	81
6	ANÁLISE E DISCUSSÃO DOS RESULTADOS	83
6.1	CARACTERIZAÇÃO DO REFERENCIAL TECNOLÓGICO PRESENTE NO UNIVERSO DAS PATENTES.....	84
6.2	CARACTERIZAÇÃO DO REFERENCIAL CIENTÍFICO PRESENTE NO CONJUNTO DAS PATENTES EM CÉLULAS-TRONCO	89

6.3 ACOPLAMENTO DOS INVENTORES EM RELAÇÃO AO REFERENCIAL CIENTÍFICO	93
6.4 ACOPLAMENTO DE PATENTES SOBRE CÉLULAS-TRONCO.....	98
6.5 ANÁLISE DE COCITAÇÃO A PARTIR DOS REFERENTES.....	104
6.6 ANÁLISE DE COCITAÇÃO DAS PATENTES SOBRE CÉLULAS-TRONCO ...	107
7 CONSIDERAÇÕES FINAIS.....	113
REFERÊNCIAS.....	119
ANEXO A – Quadro de matérias patenteáveis em biotecnologia por países	131
APÊNDICE A – Código dos países onde as patentes são depositadas.....	134
APÊNDICE B – Tabela contendo as 144 patentes analisadas com suas respectivas citações e número de artigos e patentes citadas.....	136
APÊNDICE C – Links de acoplamento entre patentes	139

1 INTRODUÇÃO

O processo de construção do conhecimento científico acontece a partir do diálogo que ocorre entre os pesquisadores que compõem a comunidade e entre eles e os documentos científicos produzidos pelos demais sujeitos que constituem a rede. Sabe-se que, para haver produção do conhecimento científico é necessário dialogar com os saberes previamente existentes. Diante disso, compreende-se que as citações realizadas, bem como a lista de referências apresentadas ao final de cada trabalho científico, refletem esse diálogo entre o pesquisador e a comunidade científica onde ele está inserido. Desse modo, observa-se, a partir da aplicação de estudos métricos, que a construção da identidade científica de um pesquisador não está atrelada apenas aos laços sociais, indicando sua capacidade de dialogar com outras áreas do conhecimento.

Dentro do cenário da Ciência da Informação, a análise da produção científica por meio do conjunto de referências que se apresentam nos artigos publicados em periódicos científicos, trabalhos em eventos, livros, dentre outros, já apresenta estabilidade. No entanto, quando investigamos tal análise no contexto dos documentos de patente, observa-se uma aplicação mais paulatina.

Embora enquanto título oficial de propriedade temporária, concedido pelo Estado ao(s) seu(s) titular(es) (BRASIL, 2015), uma patente precisa ser depositada nos países onde há interesse que o direito do(s) titular(es) à invenção seja reconhecido e protegido, passando por um período de sigilo, após sua publicação, torna-se disponível para consulta nas bases de dados especializadas ou nas bases de dados dos próprios escritórios de patentes. Pimenta (2017) considera que os documentos de patentes são importantes fontes de informações sobre o conhecimento científico e o tecnológico desenvolvidos em diversas áreas. Nesse contexto, destaca-se a importância do aprofundamento das análises sobre esse tipo de registro formal da produção científico-tecnológica.

Bufrem e Prates (2005) consideram que a área de Ciência da Informação, com seus estudos métricos, demonstra forte potencialidade na análise das relações e interações entre a Ciência, Tecnologia e Inovação (CT&I) e as demais áreas do conhecimento. Tal observação está amparada no fato de que os estudos métricos e seus subcampos, usados para analisar produção científica, já foram e seguem sendo aplicados em estudos de documentos patentes e as relações existentes entre ciência

e tecnologia. Entre os pesquisadores sobre tal abordagem metodológica, destacam-se aqui: Narin (1994), Glänzel e Meyer (2003), Leydesdorff (2008a, 2008b).

Comumente utilizada para avaliar o impacto científico das pesquisas publicadas, a análise de citação, um dos principais métodos bibliométricos, tem como premissa a relação formal existente entre o documento citante e aquele(s) citado(s). Mensura o impacto e a influência dos trabalhos científicos, a partir do reconhecimento da comunidade científica, ao inseri-los em suas listas de referências que sustentam o avanço alcançado nas pesquisas. Desse modo, entende-se a citação como um ato de reconhecimento do pesquisador citante, da importância da ciência contida no documento citado (VANZ, CAREGNATO, 2003; MACIAS-CHAPULA, 1998). Bornmann e Daniel (2008) reiteram esse entendimento e afirmam que a análise de citação é apropriada para a avaliação do desempenho científico.

Desse modo, entende-se que na Comunicação Científica, a citação tem como propósito a meritocracia e o reconhecimento dos trabalhos anteriormente publicados e que, por sua vez, fundamentam a ciência na atualidade. Contudo, no âmbito tecnológico, na análise das citações em documentos de patente, deve-se levar em conta que, além do reconhecimento, as citações podem ser decorrentes de outros motivos ou objetivos, uma vez que estas podem ser inseridas pelo inventor ou pelo examinador. Nesse sentido, citações provenientes do examinador de uma patente podem ser destinadas a refutar a sua atividade inventiva, citando outras patentes que já contemplam a inovação tecnológica, bem como para proporcionar clareza quanto ao que está sendo patenteado. Ademais, Karki (1997) considera que uma patente muito citada pelas patentes posteriores é importante e significativa na identificação das frentes tecnológicas¹.

Com base em seus estudos, Narin (1975; 1994) considera que, em função da significativa similaridade detectada entre os documentos da literatura científica e os documentos de patentes, é possível aplicar os procedimentos metodológicos bibliométricos a ambos, sob a mesma problematização. Assim como ele, Glänzel e

¹ O termo "frente tecnológica" é usado aqui por analogia ao termo "frente de pesquisa" entendido por Chen (2005) como o agrupamento emergente e transitório de conceitos e questões de pesquisa de um campo científico, definido pela delimitação resultante da citação e cocitação presente na literatura científica - uma rede em evolução de publicações científicas citadas pela frente de pesquisa. Assim, entende-se frente tecnológica como o agrupamento emergente e transitório de conceitos e questões subjacentes a um campo tecnológico, definido pela delimitação da citação e cocitação presente em seu conjunto de patentes - uma rede de patentes e conhecimento científico em evolução, citados pela frente tecnológica.

Meyer (2003) também asseguram que os estudos bibliométricos aplicados à análise das patentes seja possível e relevante, uma vez que estes são uma forma de mensurar as relações entre ciência, tecnologia e governo. Por sua vez, Leydesdorff (2008b) avalia que a análise das citações de patentes é um estudo relevante para o desenvolvimento da CT&I, uma vez que pode ser usado para mapear a base científica do conhecimento tecnológico, proteger o portfólio industrial, bem como medir o valor econômico de uma empresa ou país. Nesse cenário, destaca-se a importância e a relevante contribuição dos estudos de análise de citação para o mapeamento da base científica e tecnológica que sustentam as patentes - registros dos avanços em um determinado campo tecnológico.

As análises de citação se desdobram em dois tipos: uni-variadas e relacionais. As análises de citação uni-variadas objetivam identificar as características científicas individuais, ao passo que as relacionais se voltam para os vínculos entre esses indivíduos em uma comunidade científica (GRÁCIO, 2020). White (2001) considera que a partir da análise de citação individual (uni-variada) é possível conhecer a identidade de citação dos autores, que se desenvolve por meio da recitação sincrônica e diacrônica. Por sua vez, os estudos relacionais de citação identificam a similaridade ou proximidade entre os autores (ROSTAINING, 1996), permitindo, assim, visualizar as relações estruturais de conectividade teórico-metodológica em um campo do conhecimento. Considera-se, desse modo, que são necessárias análises uni-variadas e, também, relacionais de citação para alcançar uma compreensão mais ampla e significativa de um campo científico, bem como a identificação de domínios científicos e visualização de sua respectiva estrutura intelectual, em termos de influência (cocitação) e atores ativos (acoplamento) (GRÁCIO, 2020).

Vale ressaltar que, segundo Callon, Courtial e Penan (1995), a distinção entre análises univariadas e relacionais remete a perspectivas distintas para os estudos do desenvolvimento científico e tecnológico. Na primeira, ao quantificar a produção e o impacto, a fim de avaliar o desempenho científico e tecnológico, considera a ciência uma atividade produtiva normal, ao passo que a perspectiva relacional foca na identificação e conhecimento das relações entre autores, assim como entre tecnólogos.

De acordo com Marshakova (1981), há dois métodos de análises relacionais de citação, a saber: Acoplamento bibliográfico e Cocitação. O primeiro mede a proximidade teórico-metodológica entre dois artigos com base na intensidade do

número de documentos citados em comum. Segundo Kessler (1965), se dois artigos apresentam listas de referências semelhantes, há uma relação implícita entre eles, o que permite identificar o ambiente intelectual em que os autores atuam. Segundo Hjørland (2013), entender o acoplamento bibliográfico entre autores significa entender o grau de sobreposição da identidade de citação desses autores.

O estudo de cocitação, por sua vez, permite, por meio da análise de frequência com que dois autores ou documentos são citados juntos, averiguar os níveis de influência na área, bem como as proximidades temáticas e a interação entre os autores, a partir da perspectiva da comunidade citante (GRÁCIO; OLIVEIRA, 2013). Assim, a proximidade entre dois documentos, ou autores, não é determinada pelos autores dos documentos, mas definida pela comunidade científica que apreende seu conteúdo e estabelece conexões, durante o processo de geração de novos conhecimentos (SMALL, 1973; GRÁCIO, 2020). Essa proposta apresentada por Small (1973) foi uma resposta ao acoplamento de Kessler e anos depois, White e Griffith (1981) propuseram a versão das cocitações para autores. No entanto, foram os pesquisadores Zhao e Strotmann (2008) que propuseram o acoplamento bibliográfico aplicado a autores.

Desse modo, a comunidade científica define a estrutura do seu campo científico, a partir dos padrões de reconhecimento simultâneo da contribuição das publicações citadas concomitantemente, criando uma associação de significados (SMALL, 2004; GRÁCIO, 2020). Nesse contexto, entender os padrões de cocitação em uma literatura científica significa entender a história do reconhecimento e impacto científico das publicações (HJØRLAND, 2013).

No âmbito tecnológico, Wang, Zhang e Xu (2011) destacam a contribuição da análise de cocitação a fim de identificar as principais patentes, as semelhanças entre elas e as principais frentes tecnológicas. Rodriguez *et al.* (2015) observam que a mensuração efetiva da similaridade entre patentes em uma complexa rede de citação de patentes é uma tarefa crucial para entender o relacionamento entre elas e utilizam para esse fim os indicadores normalizados de intensidade de cocitação. Kuan e Chen (2017) apontam também que a análise de citação e o método de acoplamento bibliográfico são procedimentos válidos na detecção da similaridade entre patentes. Ainda, Zhang *et al.* (2016) destacam que as medidas de similaridade, entre elas o acoplamento bibliográfico e a cocitação, são procedimentos fundamentais para identificar relacionamentos dentro ou entre portfólios de patentes.

No entanto, embora as análises relacionais de citação, particularmente o acoplamento bibliográfico e a cocitação, possam contribuir de forma significativa para a compreensão do desenvolvimento dos diversos campos tecnológicos, estudos baseados nesses procedimentos são ainda incipientes, tendo se intensificado mais recentemente, especialmente nos últimos cinco anos. Neste sentido, segundo Castriotta e Guardo (2016), apenas alguns estudos bibliométricos tecnológicos têm utilizado o método de cocitação de patentes para o mapeamento científico e tecnológico. Além disso, nenhum estudo bibliométrico no âmbito das patentes procurou entender o impacto do conhecimento científico e tecnológico acumulado na estrutura tecnológica que dá origem aos resultados de inovação. Segundo ainda esses autores, a maioria dos estudos bibliométricos tecnológicos aborda tradicionalmente a comparação de classes e portfólios de patentes.

Hjørland e Albrechtsen (1995) mencionam que a bibliometria, como metodologia, é capaz de apresentar as tendências sociológicas no desenvolvimento do conhecimento científico, mas, segundo eles, não é capaz de interpretá-las, sendo necessário aplicar análise de domínio como aprofundamento metodológico para tal compreensão.

No âmbito da Ciência da Informação, em nível internacional, a Análise de Domínio vem sendo estudada por Hjørland e Albrechtsen (1995), Hjørland (2002, 2004, 2017), Tennis (2003, 2012) e Smiraglia (2011, 2012, 2015), por sua vez, nacionalmente temos autores como Guimarães (2014), Marteleto e Carvalho (2015), Guimarães *et al.* (2017), Amorim, Vianna e Medeiros (2018), Grácio (2018), Almeida e Dias (2019) e Evangelista, Grácio e Guimarães (2022), dentre outros, demonstrando ser um aporte para avaliação da ciência e compreensão do desenvolvimento e da socialização das diversas comunidades científicas.

Embora Hjørland e Albrechtsen (1995) não tenham apresentado uma definição precisa do que é domínio, afirmam que para se compreender a informação na Ciência da Informação, é preciso estudar os domínios do conhecimento como comunidades discursivas ou de pensamento, que por sua vez, são partes da divisão de trabalho da sociedade. Mencionam que são reflexos dos objetos de trabalho dessas comunidades: a organização do conhecimento, as formas de linguagem, as formas de comunicação, o sistema de informação, a estrutura de trabalho e os padrões de cooperação. Posteriormente, Hjørland (2002) apresenta 11 abordagens pelas quais a

Ciência da Informação pode analisar um domínio e afirma que para o fortalecimento da área, as abordagens não devem ser utilizadas isoladamente.

Sendo a bibliometria, mais especificamente a patentometria, capaz de mapear as bases científicas de patentes, entende-se que é relevante para o presente estudo identificar, a partir da análise relacional (acoplamento e cocitação), os domínios que compõe as patentes em células-tronco. Fala-se em identificar os domínios a partir da análise relacional, pois como bem menciona Mai (2005) e Tennis (2012), a nossa visão de domínio muda e o próprio domínio se reconfigura com o tempo. Além disso, Hjørland (2017) afirma que a linguagem dos domínios se distancia da linguagem geral, em virtude do aumento da especialização deles. Dessa forma, como assegura o autor, a comunicação em domínio pode ser modelada por meio de estudos métricos.

Desse modo, dada a significativa contribuição da análise de citação, uni-variada e relacional, bem como da análise de domínio para o conhecimento da base científico-tecnológica da frente tecnológica dos diferentes campos e a fim de caracterizar a proximidade das identidades tecnológicas das patentes e suas imagens ante a comunidade tecnológica em células-tronco, esta pesquisa pretende responder o seguinte problema: Qual a contribuição da análise conjunta de métodos bibliométricos uni-variados e relacionais de citação para caracterização de um domínio tecnológico formado por patentes? Diante dessa questão, tem-se como hipótese que os métodos bibliométricos uni-variados e relacionais de citação são capazes de caracterizar e identificar domínios tecnológicos no âmbito das patentes sobre células-tronco.

A fim de examinar a questão colocada, esta pesquisa objetiva analisar a contribuição dos métodos de análise uni-variadas e relacionais de citação, aplicados às patentes, para a identificação e caracterização de domínios científico-tecnológicos.

O objetivo geral desdobra-se nos seguintes objetivos específicos:

- a) Identificar o referencial tecnológico (citações a patentes) presente no universo das patentes pesquisadas;
- b) Caracterizar o referencial científico (citações à literatura não-patente), adotado no conjunto de patentes de células-tronco da pesquisa;
- c) Identificar os agrupamentos da produção tecnológica relacionadas a células-tronco, caracterizados pelas proximidades tecnológicas e científicas das patentes indexadas na base DII, evidenciados pelas

similaridades nas suas referências e mensuradas pelo acoplamento bibliográfico;

- d) Descrever os agrupamentos tecnológicos e científicos identificados a partir das imagens (citações simultâneas) das patentes em células-tronco.
- e) Reconhecer os métodos bibliométricos de acoplamento bibliográfico e cocitação entre patentes para identificação de domínios científicos-tecnológicos.

Baseado no que foi exposto, entende-se que um trabalho científico e tecnológico tem o seu reconhecimento e relevância social atestados por meio das citações e aplicabilidades em pesquisas futuras. Desta forma, compreende-se que a análise de citação (uni-variada ou relacional), enquanto abordagem metodológica, desempenha um processo social não apenas no âmbito da literatura científica, mas também tecnológica, proporcionando clareza em relação ao cenário contemporâneo. Por sua vez, desde o século XX, as células-tronco são vistas como um grande avanço da ciência e com grande potencial para o desenvolvimento da medicina regenerativa. Sendo ainda muito estudada, esbarra em questões éticas, moral e religiosa, além de sua exploração científica estar subordinada à legislação de cada país.

Ciente do papel que a célula-tronco desempenha no tratamento de doenças autoimunes e compreendendo o seu valor social e científico, entende-se que é importante analisar e conhecer o cenário não apenas a partir da literatura científica, mas também por meio do que está sendo tecnologicamente produzido. Para isso, compreende-se que é importante analisar os documentos de patentes em nível mundial. Estudos desta natureza foram desenvolvidos por Machado (2015), que analisou a cocitação de periódicos na área de células-tronco dos países que compõem o BRICS. Entretanto, analisar domínios tecnológicos de patentes sobre células-tronco concedidas em âmbito mundial, por meio de análises relacionais de citações é inédito no cenário da ciência brasileira. Contudo, ela encontra aporte teórico nas pesquisas Lai e Wu (2005), Wang, Zhang e Xu (2011) e Mogee e Koler (2005) para outras áreas tecnológicas. Lai e Wu (2005) propuseram o uso da análise de cocitação de patentes para estabelecer uma nova classificação internacional de patentes. Por sua vez, por meio da cocitação, Wang, Zhang e Xu (2011) analisaram a produção tecnológica das 500 maiores empresas norte americanas e Mogee e Koler (2005) aplicaram a análise de cocitação para identificar as frentes tecnológicas da empresa farmacêutica Eli Lilly

& Co. Ressalta-se ainda a incipiência, mesmo em âmbito internacional, da utilização do método de Acoplamento Bibliográfico para a compreensão do cenário das diversas áreas tecnológicas. Desse modo, entende-se que a utilização da análise relacional de citação, mais especificamente por meio da análise de cocitação e de acoplamento bibliográfico, contribui para a ampliação das perspectivas de entendimento e visualização do comportamento da produção tecnológica, por meio da análise objetiva e complexa dos documentos de patentes em células-tronco, ampliando a compreensão da produção tecnológica mundial nesta área.

Estudar a CT&I a partir de uma abordagem quantitativa proporciona clareza do cenário científico e tecnológico atual sobre determinada temática. Também permite averiguar onde as pesquisas estão concentradas, quais são os pesquisadores mais atuantes e quais têm maior visibilidade. Ademais, quando se trata de documentos de patentes, é possível aferir quais instituições (empresas e universidades públicas ou privadas; instituições de apoio à pesquisa ou pessoa física) realizam maior número de pedidos de patentes, em que países solicitam proteção aos seus inventos e quais áreas do conhecimento são predominantes nos pedidos de patentes. Também é possível analisar, por meio de estudos métricos, os níveis de similaridades entre os documentos, autores, assuntos, entre outras análises. Considera-se que estudos como esses são relevantes para o desenvolvimento da CT&I, tanto no âmbito nacional como internacional.

A produção científica sobre células-tronco foi elucidada por Machado (2015), em sua tese de doutorado, abordando citação e cocitação em periódicos nos países que compõem o BRICS. Quanto à produção tecnológica, o INPI (2010) realizou um estudo com abordagem quali-quantitativa com intuito de identificar como a tecnologia dessa área vem sendo protegida no Brasil. Tais estudos despertaram na autora o interesse em explorar e compreender melhor os documentos de patentes em células-tronco em âmbito global, uma vez que no Brasil, com base no artigo nove, inciso IX, da Lei de Propriedade Industrial nº 9.279, de 14 de maio de 1996, tais materiais não podem ser patenteáveis.

Nos últimos anos, notícias sobre doenças degenerativas como: câncer, Alzheimer, Parkinson, esclerose múltipla, esclerose lateral amiotrófica, artrose e glaucoma, entre outras, estão circulando facilmente nos mais variados meios de comunicação, provocando na sociedade a ansiedade de que os pesquisadores encontrem logo a cura para tais enfermidades. Motivada pelo interesse em

compreender os desenvolvimentos na área da saúde e ciente de que a medicina regenerativa está colaborando para a estabilização de muitas doenças, a autora se viu instigada a aprofundar os estudos de patentes em células-tronco. Dessa forma, fez-se necessário compreender, por meio da patentometria, como se caracterizam as pesquisas tecnológicas nessa área em âmbito global, considerando a importância do tema células-tronco para o avanço da medicina regenerativa, assunto de interesse mundial.

Wang, Zhang e Xu (2011) afirmam que estudos métricos em documentos de patentes são comumente aplicados para contagem de patentes, para medir a produtividade e o impacto nas comunidades e para analisar as classificações que mais aparecem, bem como as empresas que mais patenteiam.

Observou-se que existe uma gama significativa de estudos sobre patentes realizada em diferentes áreas do conhecimento, tais como Engenharia, Farmacologia, Química, Física entre outras (PEREIRA; FUJINO, 2014) e cada vez mais, no campo da Ciência da Informação, no qual são realizados estudos sobre patentes nas mais diversas áreas, como Biodiesel, Biotecnologia, entre outras.

Quanto à análise de citação e cocitação em patentes, destacam-se Spinak (2003), Lai e Wu (2005); Wang, Zhang e Xu, (2011) com estudos na área de microeconomia e inovação e economia industrial. Especificamente sobre a análise bibliométrica em células-tronco, temos os estudos de Cantos-Mateos e seus colaboradores (2011), que estudaram a área de células-tronco pela perspectiva das palavras-chaves, além do estudo de Zhao e Strotmann (2011), que aplicou análise de citação em pesquisa com células-tronco. Quanto aos estudos sobre patentes de células-tronco, embora Wu, Zhang e An (2012) e Cook (2012) tenham publicado trabalhos sobre o assunto, ainda se observa uma lacuna quanto à abordagem da citação em patentes de células-tronco, no Brasil e no mundo. Esse fato motivou a autora a pesquisar o assunto em âmbito mundial nos documentos indexados na Base Derwent Innovation Index, a partir do entendimento da importância dos estudos patentométricos para o desenvolvimento da CT&I.

Entende-se que, na ciência, as motivações para citar são bem diversificadas. Podem estar relacionadas tanto ao reconhecimento, bem como à proximidade temática ou teórica do autor citado com o autor que está citando. Porém, quando se trata de citação em patente, as motivações podem ser outras. Van Raan (2017) reitera que citações em publicações científicas são exclusivamente feitas pelo autor; já em

documentos de patentes, as citações podem ser feitas tanto pelo inventor, para justificar o ineditismo do invento ou qual parte do estado da arte a patente se diferencia do que já foi protegido, quanto pelo examinador, para refutar a invenção. Egger, Seliger e Woerter (2016) afirmam que as citações em patentes também são usadas para rastrear a evolução de novos campos da tecnologia, refletindo as mudanças tecnológicas e o crescimento econômico.

Por sua vez, Marco (2007) considera a citação de patente muito semelhante à citação bibliográfica, afirmando que uma inovação pode ser parcialmente baseada em uma inovação patenteada anteriormente. Wang, Zhang e Xu (2011) reiteram as afirmativas anteriores ao mencionar que as citações de patentes são indicadores do impacto de uma patente sobre outras, uma vez que as patentes citadas com maior frequência são consideradas como tendo maior influência em P & D.

Os estudos supracitados nos levam a entender que os documentos de patentes são fontes de informações pertinentes à pesquisa científica, porém, devido à sua complexidade, são pouco explorados. Com isso, escrutinar documentos de patentes sob a temática das células-tronco, adotando como método as análises métricas, torna-se um desafio instigante à estudante, visto que são objetos de valor científico, tecnológico e social.

Em 2018, a autora desta tese realizou uma pesquisa sobre a produção tecnológica em células-tronco, aplicando como metodologia a análise de citação, identificando os escritórios de patentes, investigando as maiores depositantes, as citações às patentes e à literatura não-patente (SANTOS, 2018). Nesse mesmo período, a pesquisadora Grácio (2018) defendeu sua tese de livre docência aplicando análise relacional de citação no campo dos Estudos métricos da Informação no Brasil, pesquisa que foi publicada em forma de livro posteriormente (GRÁCIO, 2020). O estudo desenvolvido por Grácio (2018) despertou um interesse pessoal em aplicar essa metodologia no cenário da tecnologia e inovação. Assim, compreendendo a importância dos estudos métricos para identificar e caracterizar domínios, essa pesquisa foi ampliada a fim de aplicar métricas que identifiquem a imagem e identidade de citação no cenário tecnológico, mais precisamente nas patentes que envolvem células-tronco e assim contribui para uma nova abordagem da análise de domínio, contemplando a comunidade tecnológica.

A presente tese está organizada em sete capítulos. Apresenta-se no primeiro capítulo a temática estudada, a definição do problema de pesquisa, hipótese, os

objetivos e a justificativa do estudo para o avanço da pesquisa no cenário tecnológico, tendo como método os Estudos Métricos da Informação.

No segundo capítulo, apresentam-se os elementos históricos e conceituais das patentes, seus subcampos, abordando os elementos que compõem os documentos de patente, a legislação em âmbito nacional e internacional, bem como as tecnologias em células-tronco.

O capítulo três dedica-se a apresentar a análise de domínio a partir dos autores que contextualizaram o conceito, bem como a sua contribuição para a presente pesquisa.

O capítulo quatro é dedicado a fundamentação teórica da análise de citação, em especial, estudos de identidade e imagem de citação, análises relacionais de citação, particularmente as análises de acoplamento bibliográfico e cocitação, buscando apresentar um panorama dos temas, relacionando às tecnologias.

No quinto capítulo, apresentam-se os procedimentos metodológicos utilizados para desenvolver as análises necessárias a fim de obter uma resposta à questão e aos objetivos inicialmente estabelecidos.

No capítulo seis, são apresentados os resultados parciais, obtidos a partir da aplicação da metodologia, respondendo os objetivos propostos desta tese.

Finalizando a pesquisa, no capítulo sete, apresentam-se parcialmente as considerações finais com um compêndio dos resultados alcançados.

2 PATENTES E INOVAÇÃO: ELEMENTOS HISTÓRICOS E CONCEITUAIS

O pensamento criativo é inerente ao ser humano. Algumas pessoas conseguem com mais facilidade externar aquilo que compõe na mente e, a partir disso, constroem objetos do nada. Como um ser pensante, o Homem sempre descobriu coisas e criou objetos que contribuíram para a sua evolução e o desenvolvimento do mundo. Ele descobriu e dominou o fogo, criou a roda e utensílios cortantes que proporcionaram a sua sobrevivência ao longo da história. Soares (1998) considera que foi na Mesopotâmia, com a criação dos primeiros centros urbanos, que a arte e tecnologia encontraram campo fértil para se expandir.

Podemos ver ao longo da história a importância do espírito inventivo para a sociedade. Benjamin Franklin, um dos líderes da Revolução Americana, inventou objetos de grande serventia social, como o aquecedor, o cateter flexível, as lentes bifocais e o para-raios. Graham Bell inventou o tão importante telefone, objeto que sofreu modificações e que é ferramenta fundamental para comunicação social. No século XV Gutenberg criou um dispositivo capaz de reproduzir palavras e textos completos por meio de caracteres. Esse dispositivo, chamado de imprensa, revolucionou a forma como as pessoas liam e reproduziam a informação. Por sua vez, a arte de imprimir foi aprimorada com a criação de um prelo a vapor, por Friedrich Koering, e assim muitos outros desenvolveram objetos que até hoje são úteis à sociedade (CAMP, 1961).

Com todo esse processo criativo, o Homem observou a necessidade de se obter alguma vantagem financeira maior sob suas invenções, uma vez que elas já indicavam que impulsionariam a economia. A primeira patente de invenção foi concedida ao engenheiro florentino Felippo Brunelleschi, em 1421, por um período de três anos pela construção de um dispositivo de transporte de mármore. No entanto, a primeira lei de patentes conhecida no mundo foi aprovada em Veneza, no ano de 1474, para proteger os artesãos que trabalhavam com cristais de Murano.

Veneza, no século XV, vivia do comércio, da compra e da venda de produtos que chegavam de navio, dessa forma, havia um cuidado com os produtos e obras de artes que lá chegavam. Os venezianos, por questões econômicas, já possuíam leis que os protegiam daqueles que quisessem criar coisas novas ou até mesmo fazer uso de criações de terceiros (CANALLI, SILVA, 2011).

A legislação de patentes de Veneza é muito semelhante às leis que conhecemos hoje. Apresenta princípios e regras que permanecem até o presente momento descritas nas leis internacionais, como por exemplo: novidade, atividade inventiva, aplicabilidade industrial, sigilo, limite de vigência do privilégio e penalidade por violação dos direitos (CANALLI, SILVA, 2011). Diante dos benefícios apresentados pela lei de Veneza, outros países observaram a necessidade de mais desenvolvimento industrial e com isso foram impulsionados a criar suas próprias leis de patentes.

Em 1623, surge na Inglaterra o *Statute of Monopolies*, que tinha como intenção coibir a concessão de privilégios e monopólios rentáveis abusivos para alguns cidadãos, promovendo maior proteção ao inventor (MACHLUP, 1958). A Revolução Industrial na Inglaterra foi o start para uma série de invenções, avanços tecnológicos, produção em grande escala e transformações das condições de vida em sociedade.

No século XVIII, ocorria a Revolução Francesa, que tinha como atores uma burguesia interessada no desenvolvimento industrial e no liberalismo econômico do país. Essa luta deu início a um sistema de propriedade industrial no país e em 1791 criou-se a Lei de Patentes da França, garantindo liberdade industrial e comercial, proporcionando caráter temporário à propriedade e ao gozo das invenções, por 5, 10 ou 15 anos (MACHLUP, 1958).

Por sua vez, os Estados Unidos (EUA) superaram o *Statute of Monopolies* e criando em 1770 o Patent Act, em que constava a exigência de descrição da invenção, a fim de que fosse aprimorada assim que caísse em domínio público. No entanto, foi o presidente George Washington quem sancionou, em 1790, a lei federal de patentes. Nela constava que a concessão da patente seria atribuída por uma comissão que determinaria o período de vigência da patente, que não poderia exceder 14 anos. Além disso, na lei constava uma taxa que deveria ser paga para tal concessão (FEDERMAN, 2006; CAMP, 1961).

No Brasil, a primeira manifestação de propriedade industrial que se tem conhecimento foi o Alvará de 28 de abril de 1809, expedido por Dom João VI, com os princípios que poderiam ser aplicáveis em pedidos de patentes. Contudo, foi Dom Pedro I que, em 28 de agosto de 1830, criou uma lei sobre concessão de patentes e foi Dom Pedro II quem regulamentou a lei, a partir da lei 3.129, de 14 de outubro de 1882 (CAMPOS; DENIG, 2011).

Todas essas leis tinham caráter territorial e, com os avanços tecnológicos e a industrialização, houve a necessidade de um tratado internacional que equilibrasse as diferenças entre elas. Assim, em 1883, diversos países, entre eles o Brasil, assinaram, na França, o primeiro acordo internacional de patentes, conhecido como Convenção da União de Paris, promovendo uma flexibilidade às legislações nacionais dos países membros, respeitando os princípios fundamentais e dando aos Estados signatários um espaço comum de direito. O acordo entrou em vigor em 1884, dando origem ao que chamamos de Sistema Internacional de Propriedade Industrial, favorecendo mais e mais o detentor da patente (FEDERMAN, 2006; CHAVES *et al.*, 2007).

Em 1893, houve uma fusão da Convenção da União de Paris e da Convenção da União de Berna (criada em 1886), dando origem ao Escritório Unificado Internacional para a Proteção da Propriedade Intelectual (BIRPI), que por sua vez criou o que hoje conhecemos como Organização Mundial de Propriedade Intelectual (OMPI) ou World Intellectual Property Organization (WIPO). Essa importante organização administra diversos acordos internacionais relativos à propriedade intelectual (PI), a fim de estimular a proteção mundial mediante cooperação entre os Estados. Além disso, cabe a ela incentivar a negociação de novos tratados, atualizar as legislações nacionais e promover a ampliação de escritórios nacionais, entre outras atividades (FEDERMAN, 2006; CHAVES *et al.*, 2007).

Em 1970, em Washington, foi firmado o Tratado de Cooperação em Matéria de Patente (PCT) a fim de facilitar a busca de potencial proteção internacional e tornar mais econômica a proteção das invenções. O tratado permite solicitar a proteção simultânea nos mais de 150 países signatários. Cabe mencionar que pedido PCT não protege a invenção em nenhum país, mas fornece ao titular da invenção um período maior para buscar direitos de patente em múltiplos países. Ao realizar o registro em um único escritório via PCT, após o prazo de 30 meses é necessário entrar na fase nacional nos países desejados, onde serão concedidos (OECD, 2009).

Embora as patentes favoreçam seus detentores e promovam diversas vantagens econômicas na indústria, os benefícios nem sempre são socialmente justos, principalmente quando não há ampla concorrência naquele determinado seguimento, permitindo que o cessionário tenha comportamento abusivo, como: impedir que surjam concorrentes para o segmento, preços abusivos e desigualdade na distribuição do produto (NUNES, 2020).

Observam-se também questões associadas às patentes de software na era digital, com muitas leis sendo revistas para tentar contemplar adequadamente as necessidades deste novo segmento. Stallman (1998) considera que tanto os desenvolvedores como os usuários de patentes de software estão vulneráveis a possíveis processos jurídicos e assegura que programas de computadores não devem ter quaisquer restrições quanto ao seu uso, ressalvados alguns direitos éticos e morais do autor.

2.1 ELEMENTOS QUE COMPÕEM AS PATENTES

A patente é um documento técnico composto por muitos elementos que podem ser objetos de análise dos Estudos Métricos da Informação. Quando se tem um conhecimento sobre o que trata cada um desses elementos é possível realizar uma coleta mais precisa das informações pertinentes à pesquisa de interesse. Conhecer a estrutura desse documento é importante tanto para o pesquisador, como para os órgãos governamentais que podem demonstrar interesse em investir recursos em determinada tecnologia, bem como para empresas que buscam investigar o mercado a partir da inovação, a fim de realizar uma prospecção tecnológica.

Donald, Kabir e Donald (2018) afirmam que muitos pesquisadores que produzem pesquisas patenteáveis, tornando o documento uma fonte rica de informações, raramente leram alguma patente. Dessa forma os autores propõem dicas para ler patentes, a fim de tornar a compreensão mais eficaz sobre as pesquisas de ponta divulgadas nas especificações de cada documento. As informações contidas nos documentos, embora pareçam um pouco complexas no primeiro momento, seguem padrões de indexação bem definidos. Algumas informações são de responsabilidade do inventor e titular, enquanto outras são de responsabilidade do escritório onde se pretende solicitar o pedido de proteção.

A OMPI fornece padrões que estruturam as informações em qualquer documento de PI, proporcionando eficiência e eficácia nos processos dos escritórios de PI em qualquer lugar do mundo. Nas patentes, cada informação é acompanhada do Código INID², recomendações relativas aos dados bibliográficos relacionados às

² INID - Internationally agreed Numbers Identification of (bibliographic for the) Data. Em português, Números acordados Internacionalmente para a Identificação de Dados. É possível localizar os códigos no documento: chrome-

patentes, criado pela WIPO (2013) a fim de facilitar a identificação dos dados. A primeira página do documento é a folha de rosto, composta por informações como: resumo, número de identificação da patente, datas (depósito, publicação, concessão), título da invenção, titular(es), inventor(es), classificação da patente, desenho da patente (Figura 1).

Entre os elementos que chamam a atenção e que identificam o documento, está o número da patente. Ele é composto de letras e números. As duas letras que aparecem no início do número, referem-se aos códigos de país, ou organização, onde o pedido de patente (ou outra invenção) foi requerido ou concedido (p.ex. GB) (Apêndice A).

A Figura 1 ilustra uma folha de rosto de patente, com seus códigos INID, os quais identificam de forma detalhada todas as informações.

Figura 1- Folha de rosto da patente US2016208243-A1

	
US 2016208243A1	
(19) United States	
(12) Patent Application Publication	(10) Pub. No.: US 2016/0208243 A1
Zhang et al.	(43) Pub. Date: Jul. 21, 2016
<hr/>	
(54) NOVEL CRISPR ENZYMES AND SYSTEMS	62/201,542, filed on Aug. 5, 2015, provisional application No. 62/205,733, filed on Aug. 16, 2015, provisional application No. 62/232,067, filed on Sep. 24, 2015.
(71) Applicants: The Broad Institute, Inc. , Cambridge, MA (US); Massachusetts Institute of Technology , Cambridge, MA (US); President and Fellows of Harvard College , Cambridge, MA (US)	Publication Classification
(72) Inventors: Feng Zhang , Cambridge, MA (US); Bernd Zetsche , Gloucester, MA (US); Ian Slaymaker , Cambridge, MA (US); Jonathan Gootenberg , Cambridge, MA (US); Omar O. Abudayyeh , Cambridge, MA (US)	(51) Int. Cl. <i>C12N 15/11</i> (2006.01) <i>C12N 15/82</i> (2006.01) <i>C12N 15/85</i> (2006.01)
(21) Appl. No.: 14/975,085	(52) U.S. Cl. CPC <i>C12N 15/11</i> (2013.01); <i>C12N 15/85</i> (2013.01); <i>C12N 15/8201</i> (2013.01)
(22) Filed: Dec. 18, 2015	(57) ABSTRACT
Related U.S. Application Data	The invention provides for systems, methods, and compositions for targeting nucleic acids. In particular, the invention provides non-naturally occurring or engineered DNA or RNA-targeting systems comprising a novel DNA or RNA-targeting CRISPR effector protein and at least one targeting nucleic acid component like a guide RNA.
(60) Provisional application No. 62/181,739, filed on Jun. 18, 2015, provisional application No. 62/193,507, filed on Jul. 16, 2015, provisional application No.	

Fonte: Pedido de patente US2016208243-A1.

Entre os códigos INID que constam da folha de rosto de uma patente destacamos (WIPO, 2013):

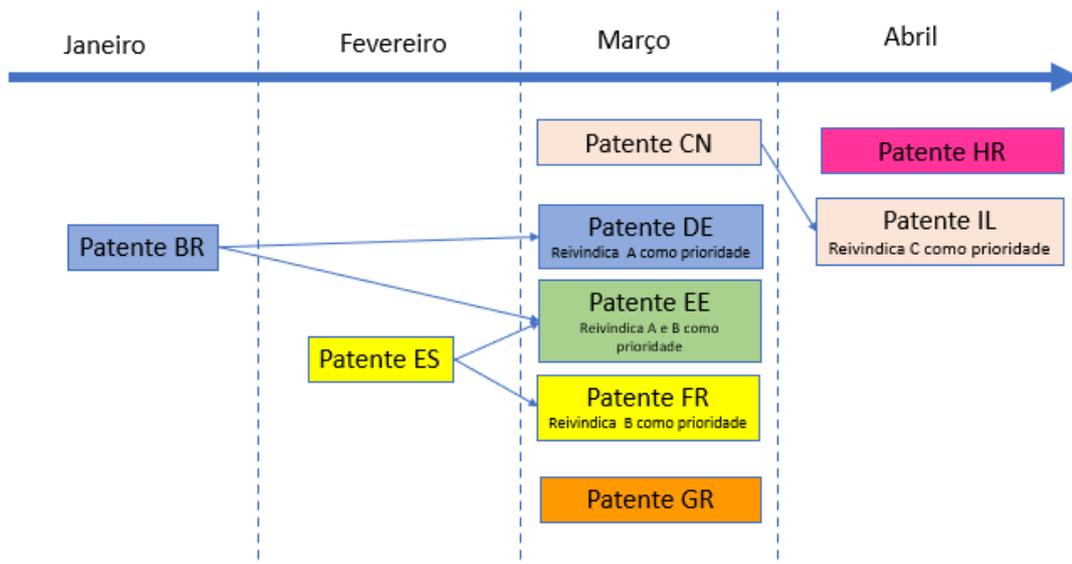
- (10) Identificação da patente
- (11) Número do documento.
- (12) Tipo de documento (Patente de invenção/Modelo de utilidade).
- (19) Nome do país ou organização regional ou internacional que publicou o documento de patente.
- (21) Número designado ao documento quando de seu depósito.
- (22) Data de depósito da solicitação.
- (31) Número designado ao primeiro depósito (prioridade do documento).
Indicação de datas.
- (32) Data de depósito da primeira solicitação (data de prioridade).
- (33) País ou países do primeiro documento (prioridade).
- (41) a (47) Referem-se a documentos de patentes publicados.
- (45) a (47) Referem-se à data de concessão da patente, geralmente é usado o número (45).
- (51) Classificação Internacional de Patente (CIP).
- (52) Classificação Nacional ou doméstica de patente.
- (54) Título da Invenção.
- (56) Lista de documentos anteriores citados pelo depositante ou encontrados pelo examinador de patentes durante a busca para exame.
- (60) Referências a outros documentos de patentes.
- (57) Resumo do conteúdo do documento.
- (71) Nome do depositante.
- (72) Nome do inventor, se conhecido.
- (73) Nome de quem detém os direitos sobre a patente.
- (74) Nome do procurador ou agente.
- (75) Nome do inventor, quando também for o depositante.

Qualquer documento de patente, independente do escritório onde foi solicitado o pedido de patente, segue esse padrão. Cabe salientar que a escolha de um determinado escritório para proteção prioritária da invenção pode indicar potencial tecnológico bem desenvolvido no país. Desta forma, a análise geográfica realizada a partir dos escritórios torna-se um indicador interessante para compreender como se configuram as proteções de determinada tecnologia em âmbito mundial.

Os escritórios podem ser divididos em regional, nacional e internacional. Uma mesma patente pode ser protegida em qualquer um deles ou em todos, desde que a invenção atenda a respectiva legislação. Sempre que uma patente é solicitada, recebe um número de identificação contendo código do escritório/país (US), o ano da publicação (2016), o número de série (208243) e o código de status que indica o tipo de documento ou estágio da publicação (A1) (Figura 1). Para esta análise, é utilizado o código do país contido no campo PN (*priority number*) tanto das patentes prioritárias como nas famílias de patentes³ deste estudo.

A família de patente é composta por várias patentes ou pedidos de patentes que estão relacionados entre si pela patente prioritária⁴ (Figura 2). Esta refere-se à primeira patente, ou seja, a mais antiga. Conforme ilustrado no Manual Básico para Proteção por Patentes Invenções, Modelos de Utilidade e Certificados de Adição (MINISTÉRIO DA ECONOMIA, 2021).

Figura 2 – Exemplo de família de patentes



Fonte: adaptado do Ministério da Economia (2021).

Conforme observa-se na Figura 2, há três famílias de patentes ilustradas. A primeira família é composta pelas patentes BR (Brasil) – sendo ela prioritária - patente

³ [chrome-extension://efaidnbnmnibpcjpcglclefindmkaj/https://www.gov.br/inpi/pt-br/servicos/patentes/guia-basico/ManualdePatentes20210706.pdf](https://www.gov.br/inpi/pt-br/servicos/patentes/guia-basico/ManualdePatentes20210706.pdf)

⁴ É a primeira patente ou pedido de patente apresentado dentro da família, ou seja, o pedido mais antigo.

DE (Alemanha) e documento EE (Estônia). A segunda família também é composta por três documentos: patente ES (Espanha - prioritária), patente EE (Estônia) e patente FR (França). A terceira família é composta da patente CN (China - prioritária) e IL (Israel). As patentes GR (Grécia) e HR (Croácia) não apresentam família, visto que não estão vinculados a nenhuma patente prioritária, mas podem se tornar patentes prioritárias, se houver um pedido de reivindicação prioritário para elas. Por sua vez, as patentes BR (Brasil), ES (Espanha) e CN (China) não participam da mesma patente.

A Classificação Internacional de Patentes (CIP)⁵, código 51 do INID, foi criada em 1971 a partir do Acordo de Estrasburgo e passou a vigorar em 1975. Embora seja a classificação universal dos escritórios de patentes, sabe-se que alguns escritórios como o USPTO e bases como a própria DII possuem suas próprias classificações, o que não anula o uso dela. A CIP auxilia na busca e recuperação de documentos de patentes, organiza os documentos de acordo com as informações tecnológicas e facilita a investigação do estado da técnica. Todos os pedidos de patentes são categorizados conforme as áreas tecnológicas. Possui 8 grandes seções, indicadas pelas letras A, B, C, D, E, F, G e H, com cerca de 70 mil subdivisões. As seções são:

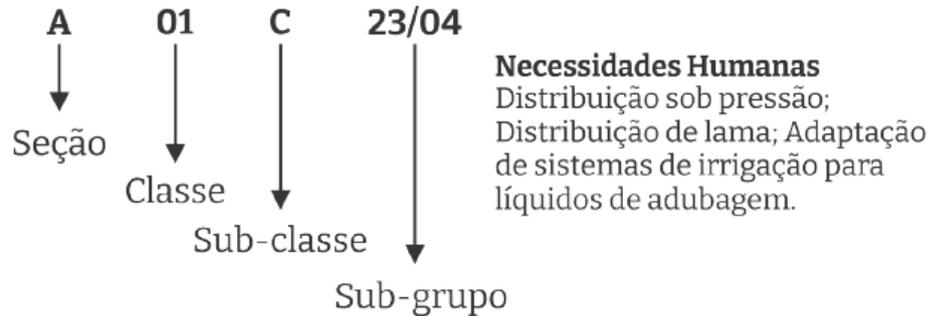
- (A) Necessidades Humanas.
- (B) Operações de Processamento; Transporte.
- (C) Química e Metalurgia.
- (D) Têxteis e Papel.
- (E) Construções Fixas.
- (F) Engenharia Mecânica, Iluminação; Aquecimento; Armas; Explosão.
- (G) Física.
- (H) Eletricidade.

Após a seção indicada por uma das oito letras, a classificação é seguida de um número de até dois dígitos indicando a classe, sendo algumas vezes complementado com letra que indica a subclasse, um número que varia de 1-3 dígitos indicando o

⁵ Acesso à Classificação Internacional de Patentes.
(<https://ipcpub.wipo.int/?notion=scheme&version=20220101&symbol=none&menulang=en&lang=en&viewmode=f&fipopc=no&showdeleted=yes&indexes=no&headings=yes¬es=yes&direction=02n&initial=A&cwid=none&tree=no&searchmode=smart>)

grupo principal e mais uma barra oblíqua "/" e números contendo até três dígitos indicando o subgrupo (Figura 3).

Figura 3 – Estrutura da Classificação Internacional de Patente



Fonte: Elaborado por Bochi, Gabriel Júnior e Moura (2021).

Uma patente pode apresentar diversas classificações e elas são indispensáveis para a recuperação de documentos de patentes durante a busca pela anterioridade. Esta recuperação é utilizada por autoridades concedentes de patentes, potenciais inventores, unidades de pesquisa e desenvolvimento e outras partes interessadas na aplicação ou desenvolvimento da tecnologia.

Para manter-se atualizada a CIP por um processo de revisão contínuo e uma nova versão é divulgada anualmente.

2.2 LEGISLAÇÕES

Diante desse breve resgate histórico, além da explicação sobre a estrutura dos documentos de patente, é possível observar o papel fundamental que a patente tem para o desenvolvimento econômico e tecnológico de um país. Estados com maior potencial tecnológico são, na sua maioria, aqueles que grandes empresas e instituições apresentam maior disposição em solicitar patente. Sendo assim, cabe evocar que as patentes, uma vez que são concessões feitas ao inventor, precisam ser depositadas nos países onde há interesse em proteger a invenção e, desse modo, devem atender a legislação vigente de cada Estado (BRASIL, 1996).

No Brasil, a lei de Propriedade Industrial de 14 de maio de 1996 garante, no artigo 6º, que o autor da invenção tem o direito de obter patente sobre sua criação, desde que a invenção atenda os critérios de patenteabilidade (BRASIL, 1996).

Como afirmam Bochi, Gabriel Júnior e Moura (2020), a Lei nº 9.279 de 14 de maio de 1996 de propriedade industrial apresenta critérios direcionados aos inventores e à atividade industrial e indicação de que nem tudo o que é criado recebe proteção, mas apenas aquelas que atendem os critérios legalmente estabelecidos. Coelho (1997, p.73) afirma que a patente se refere à invenção, definida como “[...] o ato original do gênio humano [...]”, atendendo os requisitos da novidade, da atividade inventiva, da aplicação industrial e do não-impedimento.

Dessa forma, vale ressaltar que a patente é um documento, concedido pelo Estado, que confere o direito de propriedade e de uso exclusivo de uma invenção ao seu requerente, que por sua vez pode ser uma pessoa física ou jurídica. A solicitação da patente passa por todo um processo de análise até que seja concedido o documento. Embora esse processo possa levar anos dependendo do país de análise, sua proteção legal já é considerada a partir do depósito (INPI, 2020).

Salienta-se que as leis de cada país apresentam critérios que são relevantes para aquele Estado e que por sua vez garantem a sua segurança. Como exemplo, cita-se o Japão, que apresentou na Japan Patent Act, publicada pela Lei nº 140 de 13 de abril de 1959, cláusulas sobre a violação de patentes, uma definição de invenção e licenciamento exclusivo (JAPÃO, 2010). Já a Lei de Patentes e Desenhos Indianos, criada em 1911, foi elaborada com o objetivo de favorecer os interesses das empresas da metrópole, garantindo seu controle do mercado indiano (INPI, 2007).

Destaca-se que essas leis são elaboradas para assegurar proteção e benefícios financeiros ao Estado, as empresas, aos órgãos governamentais e demais instituições. Para dirimir as burocracias impostas pelas diversas leis, amenizar os custos de depósito em vários países e, até mesmo, ter acesso às informações técnicas relativas à invenção e ao depósito da patente em determinado país, criou-se o Tratado de Cooperação de Patentes (PCT), administrado pela WIPO.

Esse tratado tem um efeito jurídico multinacional, que auxilia os inventores na busca de potencial proteção internacional de patentes para seus inventos, permitindo adiar por mais tempo o pagamento das principais taxas e tomar decisões mais precisas que atendam às necessidades da sua instituição, visto que ganha mais 18 meses, a contar da data de depósito, para definir onde quer reivindicar a proteção de sua patente. Além disso, mediante a apresentação de um pedido de patente internacional sob o PCT, os candidatos podem procurar simultaneamente a proteção

de uma invenção na maioria dos países do mundo que pertencem ao tratado, depositando um único pedido de patente internacional (WIPO, [2020]).

Observa-se que para se chegar à concessão de uma patente, passa-se por muitas etapas legais e quando se trata de patentes de organismos vivos outras análises precisam ser feitas, como por exemplo estudar e compreender as leis de biotecnologia vigente em cada Estado. Essas discussões são fundamentais para compreender as leis concernentes ao material genético. Conforme apresentado na dissertação de Santos (2018), trabalho base para essa pesquisa juntamente com a tese de livre docência de Grácio (2020), dentre os produtos capazes de ser patenteados, encontram-se os biotecnológicos, que encontram na legislação, principalmente no Brasil, algumas restrições quanto ao patenteamento de organismos humanos.

2.3 TECNOLOGIA EM CÉLULAS-TRONCO

Pesquisas com células-tronco passam por constantes discussões de caráter moral, ético e religioso. Assim, como visto anteriormente, cada país estabelece leis específicas para atividades que utilizam organismo humano a fins de pesquisa.

Diniz e Avelino (2009) trazem um panorama das pesquisas com células-tronco em âmbito internacional, em que apresentam o comportamento de 25 países em relação à pesquisa com material genético. Desses países, a Alemanha é a única nação que permite a pesquisa embrionária desde que seja feita com linhagens importadas, ou seja, os pesquisadores alemães não podem fazer uso de material genético nativo para o desenvolvimento de seus estudos. Ainda no estudo de Diniz e Avelino, podemos observar que 23 países são adeptos, não sem suas ressalvas legais, ao uso de embriões nacionais e importados. Por outro lado, a Itália, país com um forte potencial tecnológico para esse tipo de pesquisa, não permite pesquisa embrionária. Dentre os países que autorizam pesquisa com células-tronco embrionárias, os EUA é o país de maior destaque em publicação de pesquisa experimental, no entanto sua lei não autoriza que estudos sejam financiados com recursos federais.

Mendez-Otero e seus colaboradores (2007, p. 5) corroboram com o estudo ao trazerem dados que comprovam a importância das patentes em células-tronco. Segundo eles, nos EUA, uma vez que a legislação permite o patenteamento de “[...]”

linhagens celulares e tecnológicas que envolvam o uso de células-tronco [...]”, mais de 2000 patentes foram concedidas para uso e tratamento de doenças. Os autores afirmam que o cenário mercadológico envolvendo terapias celulares gerava cerca de US\$24,6 bilhões em 2005 e pode ser três vezes maior em 2010, chegando a aproximadamente US\$70,00 bilhões. Além disso, segundo eles, existem mais de cem corporações que detêm patentes voltadas ao uso e manuseio de terapia com células-tronco. Cabe observar que esse país tem uma vasta trajetória em estudos relacionados a biotecnologia.

Rocha (2008) afirma que o Projeto Genoma Humano (PGH), financiado pelo governo americano e com a parceria de diversos países, tornou evidente a possibilidade de retorno financeiro vindo da decodificação do genoma humano, contando também com a colaboração financeira de grandes empresas privadas. Em virtude do alto investimento nessa pesquisa e visando o forte retorno financeiro que ela traria ao país e às grandes empresas investidoras, foi necessário garantir esse retorno estabelecendo leis de patente de organismos vivos.

No Brasil, a Lei de Propriedade Industrial nº 9.279, de 14 de maio de 1996, é que regula direitos e obrigações relativos à propriedade industrial, e nela consta que as células obtidas diretamente de um ser vivo não são passíveis de patenteabilidade conforme os art. 10 (IX) e 18 (III). Porém, os processos que envolvem as células são patenteáveis, desde que não incidam no art. 18 (I) que diz: “Não é patenteável [...] o que for contrário a moral, aos bons costumes e à segurança, à ordem e à saúde públicas” e atendam aos três requisitos de patenteamento – novidade, atividade inventiva e aplicação industrial (BRASIL, 1996, grifo nosso).

A lei brasileira de Biossegurança nº 11.105, de 24 de março de 2005, atualizou os termos da regulação de organismos geneticamente modificados. O artigo 5º da lei assegura, com restrição, a manipulação de embriões humanos, produzidos por fertilização in vitro, para coleta de células-tronco. Contudo, algumas discussões foram levantadas sobre o uso desse material, alegando que a vida acontece mediante a fecundação e que o uso desses embriões são um atentado à vida e à dignidade da pessoa humana.

O Instituto Nacional de Propriedade Industrial (INPI) publicou a Resolução 144, que institui as Diretrizes de Exames de Pedidos de Patente na área de Biotecnologia (BRASIL, 2015), auxiliando na compreensão do que pode e não pode ser patenteado nessa área. Conforme a resolução, em conformidade com a Lei de Propriedade

Industrial, as células obtidas diretamente de um animal não podem ser patenteadas, porém as composições contendo essas células, os processos de obtenção de célula-tronco, bem como o seu uso, são operações que podem ser patenteadas.

Entende-se que as disparidades legais de cada país dificultam a compreensão do que é aceitável no patenteamento de células-tronco, uma vez que cada Estado tem seus critérios de aceitabilidade para patenteamento. No entanto, Zucoloto e Freitas (2013) conseguiram elucidar as legislações sobre materiais patenteáveis em Biotecnologia de países como Estados Unidos, Japão, Índia, China, União Europeia e Brasil (Anexo A).

3 ANÁLISE DE DOMÍNIO

A construção do conhecimento científico e tecnológico se dá a partir das relações sociais e de trabalho de atores de uma determinada comunidade discursiva, materializado em registros documentais, que abrangem os artigos científicos, os livros, as comunicações em eventos, as patentes, entre outros. A socialização dessas produções proporciona um avanço em forma helicoidal no conhecimento, mantendo-se em constante desenvolvimento (GUIMARÃES, 2017). Ciente disso, analisar a produção científica e tecnológica é fundamental para identificar o comportamento da ciência. Castanha, Rosa e Grácio (2016) consideram que os estudos bibliométricos constituem uma abordagem científica que fornece informações importantes sobre um domínio. Nesse contexto, salienta-se a importância de conceituar domínio.

O termo Análise de Domínio foi aplicado pela primeira vez em uma tese na área de Ciência da Computação, em que Neighbors (1981) tratava sobre análise de domínio no uso correto dos componentes de software. Ele definiu o termo como uma tentativa de identificar objetos e relações entre aquilo que os indivíduos que atuam no domínio consideram importante. A partir desse conceito, a análise de domínio passou a ser o amparo teórico para avaliar o cenário da produtividade científica em diferentes perspectivas.

O conceito de domínio pode ser compreendido a partir de distintos olhares, conforme a perspectiva do indivíduo que compreende aquele domínio. Mai (2005, p.605), por exemplo, considera que domínio pode ser uma “[...] área de especialidade, um conjunto literário ou um grupo de pessoas trabalhando juntas em uma organização [...]”. Assim, para o autor, o cerne analítico de domínio está em estudar as atividades e seus produtos para obter percepções sobre estruturas e significados daquilo que existe. A suposição é que os domínios produzem artefatos que podem ser usados para estudá-los e compreendê-los, permitindo que decisões sejam tomadas adequadamente.

Na Ciência da Informação, o conceito de domínio foi introduzido por Hjørland e Albrechtsen (1995) e argumentaram que a Análise de Domínio é um importante método para avaliação da ciência, uma vez que permite investigar o processo evolutivo das comunidades e suas relações. Neste contexto, segundo eles, um domínio científico, é fruto de uma comunidade discursiva e do papel desempenhado por ela. Para eles, a organização do conhecimento, a estrutura de trabalho, as

estratégias de cooperação, a linguagem e os meios de comunicação, os sistemas de informação e os critérios de relevância são a representação dos objetos de trabalho dessas comunidades e de seu papel na sociedade. Mai (2005) afirma que um domínio é um conceito aberto e em evolução que se desenvolve à medida que o conceito for usado e aplicado na pesquisa e na prática.

Beghtol (1995) apontou que estudos de análise de domínio são muito relevantes para a ciência e prospectou a ideia de que esses trabalhos poderiam evoluir de bases teóricas para métodos quantitativos, contribuindo para o desenvolvimento de vários tipos de linguagem de documentação em diversos domínios.

Castanha, Rosas e Grácio (2016) consideram que a análise de domínio busca a integração entre o contexto individual e social das comunidades, onde estão inseridas e os conceitos de informação ganham sentido quando ocorre o compartilhamento entre essas diferentes comunidades e seus membros. Sob essas perspectivas, Hjørland (2002) considera que a estrutura dos estudos de análise de domínio está fundamentada em 11 abordagens, que são:

- 1) **Produção de guias de literatura:** corresponde às publicações que listam e descrevem o sistema de recursos da informação em uma ou mais áreas. É uma espécie de bibliografia de documentos num domínio, mas se difere das bibliografias típicas sobre assuntos.
- 2) **Elaboração de classificações especiais e tesouros:** os tesouros são vocabulários controlados de um domínio específico, e a metodologia de seu planejamento pode também ser considerada como uma forma de análise de domínio.
- 3) **Indexação e recuperação da informação:** a pesquisa sobre indexação, representação e recuperação de documentos deve ser capaz de avaliar as práticas pouco desenvolvidas e auxiliar nas questões de recuperação.
- 4) **Estudos empíricos de usuários:** é um meio para se tentar medir a necessidade de informação dos usuários estudando ou questionando seu comportamento, cuja resposta nem sempre é confiável, segundo o autor. Em muitos casos, os usuários não sabem qual documento procurar e, quando encontram a informação, pode acontecer de não reconhecerem como a informação buscada.

- 5) **Estudos bibliométricos:** abordagem consistente para analisar e caracterizar um domínio científico que, quando aplicado com outra abordagem qualitativa, enriquece a análise e a compreensão de um domínio.
- 6) **Estudos históricos:** trata-se de compreender documentos, organizações, sistemas, conhecimento e informação, numa perspectiva histórica. Os métodos históricos são frequentemente capazes de fornecer uma perspectiva mais profunda e mais coerente quando comparados com outros tipos de pesquisa de caráter não histórico. O conhecimento de fatos anteriores favorece a ampliação e aprofundamento das questões.
- 7) **Estudos de documentos e estilos:** são disciplinas ou comunidades de discurso diferentes que desenvolvem tipos especiais de documentos adaptados às suas necessidades específicas.
- 8) **Estudos epistemológicos e críticos:** as teorias da epistemologia são as mais fundamentais teorias de relevância e qualquer questão teórica em ciência da informação é baseada em suposições epistemológicas.
- 9) **Estudos terminológicos, linguagens para propósitos específicos (LSP), semântica de bases de dados e estudos de discurso:** a linguagem e a terminologia são objetos muito importantes para a Ciência da Informação porque afetam o pensamento e, portanto, as questões que se colocam nos bancos de dados, bem como os textos que se buscam.
- 10) **Estruturas e instituições da comunicação científica:** o estudo das estruturas da divisão interna do trabalho dentro dos domínios e a troca de informação entre os domínios fornece informação útil para a compreensão da função de tipos específicos de serviços de documentos.
- 11) **Cognição científica, conhecimento especializado e inteligência artificial:** o objetivo central da ciência cognitiva na Ciência da Informação é fornecer aos usuários informações que possam ajudá-los a avaliar a validade das diferentes alegações de conhecimento e a estabelecer sua própria visão em alguma questão baseada nos estudos de todos os argumentos disponíveis.

Hjørland (2002) considera que os resultados são mais favoráveis quando se combinam diferentes abordagens. Para esta pesquisa, considera-se que os estudos bibliométricos, mais precisamente Análise de Citação e de Acoplamento Bibliográfico, constituem método significativo para analisar e caracterizar um domínio

tecnológico, que combinados com a abordagem epistemológica e crítica proporcionam contextualização para o estudo de análise de domínio.

Castanha, Rosas e Grácio (2016) consideram ainda que a partir da análise de domínio, é possível avaliar o que é realmente importante em um campo científico, para que aspectos como tendências, padrões, processos, pensamentos dominantes, agentes e suas relações possam ser identificados e analisados, tornando-se uma abordagem significativa para a caracterização e a avaliação da ciência.

Nesse contexto, Tennis (2003) observa que é necessário reconhecer o cerne de determinado conhecimento para analisar um domínio, independentemente de sua magnitude. Ele direciona seus estudos sobre domínio para questões metodológicas, em lugar das conceituais, buscando complementar as abordagens definidas por Hjørland (2002). Considera que são necessários dois eixos para definir os parâmetros de um domínio, o primeiro eixo estabelece parâmetros sobre os nomes e a extensão do domínio, que é seu escopo total, denominado área de modulação; já o segundo eixo, denominado grau de especialização, qualifica e estabelece a intensão de um domínio.

Ainda, como bem observado por Pavanelli (2012), os estudos bibliométricos sob o escopo da análise de domínio são importantes, pois identificam as conexões detalhadas e reais entre documentos individuais, representando os conhecimentos dos pesquisadores a partir dos artigos, patentes e demais. Observando a crescente proximidade entre a produção científica e tecnológica, considera-se que estudos no universo das patentes baseados no método de análise das citações, uma vez que estas definem um domínio (SMIRAGLIA, 2011), são fundamentais para revelar com veracidade o desenvolvimento e as mudanças da estrutura tecnológica estabelecida em um domínio.

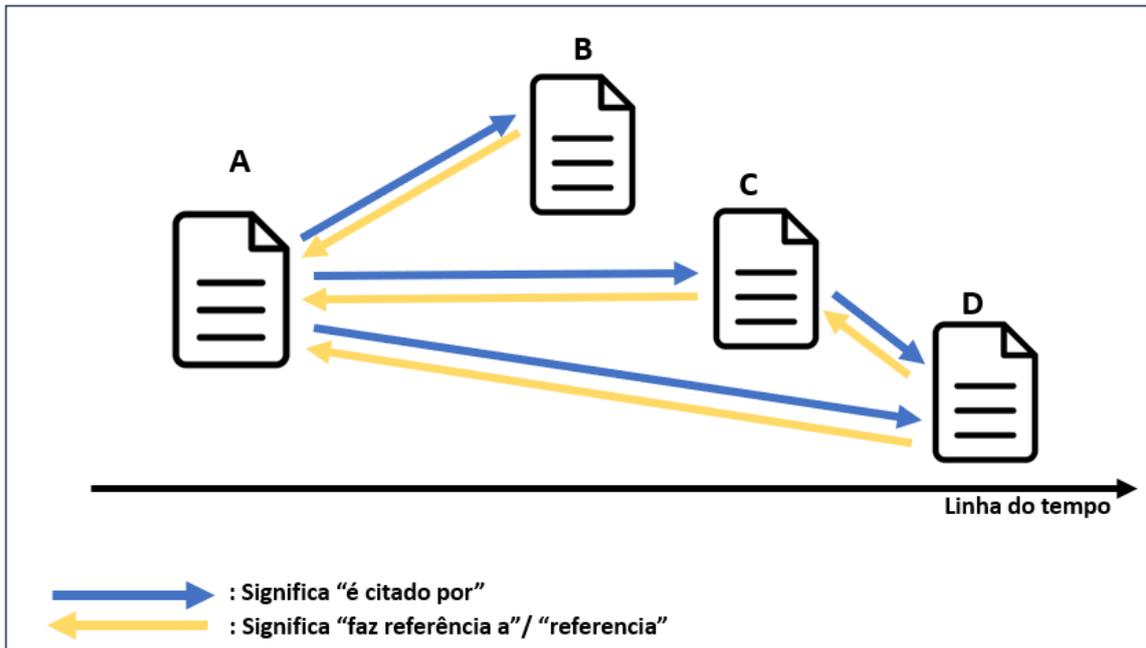
4 ANÁLISE DE CITAÇÃO

Uma das estratégias utilizadas para analisar epistemologicamente as publicações científicas é examinar o diálogo entre seus autores e aqueles nelas citados. Assim, a lista de referência apresentada no final da publicação diz muito sobre o desenvolvimento científico do autor, assim como o seu meio profissional e personalidade, visto que nela conseguimos observar o ambiente teórico em que ele fundamenta seus estudos, os autores que regem sua trajetória acadêmica e aqueles que dialogam com suas ideias, trazendo contrapontos para os seus resultados. Esse diálogo chamamos de citação.

A citação é fundamental na construção do conhecimento científico e exige, desde o século XIX, que seus locutores referenciem os pensadores que influenciam suas publicações. Smith (1981) considera que estudar as referências que acompanham cada documento é prática crucial para o bom desempenho da pesquisa científica, visto que essas referências apontam para estudos anteriores que são, em parte, base para a ciência atual. A autora afirma ainda que a citação representa uma relação entre quem cita e quem foi citado e que a natureza dessa relação é algo complexo de ser caracterizado. Garfield (1972) e Macias-Chapula (1998) apontam que há diversos motivos que levam um autor a citar, entre eles: dar crédito a trabalhos importantes e fazer críticas aos trabalhos anteriores. Além disso, citar concede credibilidade a pesquisa, validando os resultados e fundamentando o estudo com base em seus antecessores.

Contudo, é importante dizer que citação e referência, no contexto da análise de citação, são termos distintos. Egghe e Rousseau (1990) afirmam que referência é o reconhecimento que um documento dá a outro documento e citação é o reconhecimento que um documento recebe de outro. Assim, para esses autores, referência é um conceito retrospectivo (sincrônico) e citação é um conceito prospectivo (diacrônico), em que um olha para o passado e o outro para o futuro, respectivamente, conforme elucidado na Figura 4.

Figura 4 – Relação de citação e referência entre os documentos A, B, C e D



Fonte: adaptado da figura III.1.1 de Egghe e Rousseau (1990. P. 204).

Conforme podemos ver na linha do tempo da Figura 4, B, C e D fazem referência à pesquisa A, ou seja, a direção do vetor (amarelo) é para trás (sincrônico). Por outro lado, A é citado por B, C e D, caracterizado pelo vetor (azul) com direção para frente na linha do tempo (diacrônico). Além disso, C e D citam-se reciprocamente.

A análise de citação traz à luz da ciência o impacto científico que alguns trabalhos exercem sobre outros e permite observar a curva de citação de um trabalho, o que Egghe e Rousseau (2000) chamam esses processos de envelhecimento da ciência, ou seja, após a publicação, a curva de aumento no número de citações recebidas pelo artigo indicam seu estado de envelhecimento ou maturação.

A frequência com que um autor é citado e o impacto que provoca na comunicação científica indica sua influência e a visibilidade dentro da comunidade que está inserido e isso pode estar relacionado a diversos fatores como relevância, validade, relações sociais entre citante e citado, premiações e honrarias como Nobel e assim por diante (GARFIELD, 1972; BORNMANN; DANIEL, 2008; VANZ, 2004).

Observa-se que a citação está muito relacionada a meritocracia e reconhecimento dos trabalhos publicados anteriormente, ou seja, não é apenas uma prática metodológica que traz embasamento teórico ao trabalho, visto que há uma coerência no ato de citar.

No entanto, quando analisamos documentos de patente a prática de citação é composta de outros elementos que não somente o reconhecimento. No processo de solicitação de uma patente, o estado da arte e o estado da técnica são dois aspectos que devem ser levados em consideração. O estado da arte é um mapeamento de toda a produção científica de um determinado assunto. Essa atividade é uma prática importante, pois permite verificar se a solicitação da patente atende os critérios de patenteabilidade (novidade, atividade inventiva e a aplicação industrial), respeitando o que está descrito no parágrafo 1º do Art. 11 da Lei de Propriedade Industrial:

A invenção e o modelo de utilidade são considerados novos quando não compreendidos no estado da técnica. O estado da técnica é constituído por tudo aquilo tornado acessível ao público antes da data de depósito do pedido de patente, por descrição escrita ou oral, por uso ou qualquer outro meio, no Brasil ou no exterior, ressalvado o disposto nos arts. 12, 16 e 17.” (BRASIL, 1996, grifo nosso).

Sendo assim, se o indivíduo divulga sua invenção ou modelo de utilidade antes de entrar com o pedido de patente, todo o conteúdo publicado é considerado estado da técnica, não se encaixando no critério novidade e a citação contribui para identificar se tais critérios estão sendo contemplados ou não.

Pereira e Fujino (2014) afirmam que as informações apresentadas pelo inventor da patente no documento depositado, permitem que o examinador identifique se há ineditismos no invento ou é apenas uma cópia de patentes anteriores. Além disso, faz parte da atividade do examinador da patente citar outras patentes ou demais documentos com o intuito de refutar ou aprovar a atividade inventiva, proporcionado mais clareza no que está sendo patenteado. Nesse contexto, uma patente muito citada pelas patentes posteriores pode indicar que é de grande relevância, sendo usada para identificação de áreas tecnológicas de ponta (KARKI, 1997). Reiterando essas ideias, Marco (2007) aponta que as citações de patentes sinalizam a qualidade das patentes, medem o fluxo de conhecimento tecnológico e permitem investigar o comportamento estratégico das empresas e instituições.

Mattson (2011) aponta que, no desenvolvimento de um processo de pedido de patente, as citações podem ser adicionadas tanto pelo inventor, quanto pelo requerente, que pode não ser o próprio inventor, uma vez que pode ser a instituição que detém os direitos de propriedade. Essas citações são aferidas pelo examinador que determina a novidade do invento, aceitando-o ou recusando-o, pela comparação

da invenção com o estado da arte e pode incluir, se assim considerar necessário, outras referências.

Um estudo realizado por Camargo (2011), em que foram analisados 3045 trabalhos acadêmicos (dissertações e teses) dos alunos de engenharia da Universidade de Campinas (Unicamp), observou-se que as patentes eram a principal ou uma das mais importantes fontes de informação consultadas por eles. No estudo, a autora concluiu que na maioria dos trabalhos analisados, a citação de documentos de patentes não chegou a 5%, indicando que ainda não havia um reconhecimento dos documentos de patentes como importantes fontes de pesquisa. Todavia, Martínez-Martínez-Méndez, Pastor-Sánchez e López-Carreño (2010) afirmam que as universidades espanholas raramente reconhecem um documento de patente como indicador relevante para o desenvolvimento da pesquisa, concentrando-se mais nas publicações em periódicos com fator de impacto. Esses resultados contrastam com o Manual de Estatística de Patentes da OCDE (2009), que assegura que as patentes são um dos indicadores mais importantes para medir a produtividade científica e tecnológica de um país.

Compreendendo a importância das informações contidas nas patentes para medir o desempenho econômico que um país, Quoniam, Kniess e Mazieri (2014) afirmam que é preciso entender que esses documentos são base de indicadores de produção, um dos indicadores cientométricos pensados por Narin (1976), e que trazem resultados muito importantes para a CT&I. Conforme esse autor, a maioria das pesquisas científicas está voltada para a literatura científica e para as técnicas bibliométricas aplicadas à ciência e aos artigos científicos, sendo aplicáveis à tecnologia e às patentes de forma mais tangencial. Tendo essa ideia como base, o autor buscou, a partir dos estudos de Cole e Eales (1917), desenvolver técnicas bibliométricas que sejam aplicadas aos documentos de patentes e citações de patentes, bem como na avaliação de atividades tecnológicas.

Nesse contexto, Lucas, Garcia-Zorita e Sanz-Casado (2013) lembram que a análise métrica aplicada aos documentos de patente permite identificar o nível econômico e de inovação de determinada nação e/ou setor industrial, além de proporcionar uma melhor compreensão sobre o conhecimento científico que se transforma em conhecimento tecnológico, bem como a capacidade competitiva de empresas, universidades e países.

No entanto, Van Raan (2017) considera que estudos de citações de patentes, tanto de patentes como de literatura científica, não é algo simples. O autor afirma que as informações relativas à literatura não patente (científica) muitas vezes são incompletas e não unificadas, podendo conter várias referências distintas que apontam para a mesma publicação científica. Além disso, Van Raan (2017) lembra que os examinadores desempenham um papel importante na adição de citações às patentes, mas afirma que as referências fornecidas pelo examinador, no entanto, não são isentas de problemas, apresentando alguns efeitos negativos no que tange distância geográfica na probabilidade de capturar patentes anteriores.

Nesse cenário, a análise de citação de patentes pode ser usada para entender a difusão da tecnologia. Alinhados a esse entendimento, Trajtenberg, Henderson e Jaffe (1997) apontaram que os estudos de redes de patentes contribuem para melhor estabelecer as relações entre patentes, uma vez que elas apresentam uma revisão do estado da técnica, evidenciando a relevância da contribuição de estudos de citação para compreensão desse universo tecnológico.

O uso de dados de citação de patentes pode ser vantajoso em termos de validade e confiabilidade, visto que para receber uma patente, um inventor deve explorar cuidadosamente a anterioridade. Acima de tudo, um examinador deve avaliar tanto a invenção quanto o escopo das reivindicações solicitadas, realizando uma análise aprofundada das patentes existentes. Meyer (2000) concluiu, nesse sentido, que é fundamental obter uma compreensão completa das práticas de patenteamento para interpretar corretamente os dados de citações de patentes.

4.1 INDICADORES METRICOS

No cenário tecnológico, os indivíduos que contribuem para a progressão socioeconômica querem entender as fontes e mecanismos pelos quais novas descobertas e inovações são desenvolvidas e quais são as consequências sociais dessas criações. Para isso, o uso de indicadores é fundamental, uma vez que mensuram, dão visibilidade e contribuem para tomada de decisão no que se refere a investimentos em ciência, tecnologia e inovação.

Os indicadores bibliométricos, definidos por Pritchard (1969) como a aplicação de métodos estatísticos a obras como livros, desempenha um papel fundamental nesse contexto. Eles são aplicados para analisar e medir as atividades científicas e

tecnológicos, tanto no ambiente acadêmico-científico, como empresarial e governamental, em diversos níveis.

Narin (1976) categorizou os indicadores em produção e citação, em que ambos são produtos da ciência e da tecnologia. Esses indicadores bibliométricos, posteriormente, receberam mais uma categoria, incluindo-se assim o indicador de ligação, voltado às análises de citação, concomitante, entre artigos e patentes e vice e versa.

Os indicadores de ligação são os métodos mais eficazes de mensurar a produção científica e tecnológica, visto que permitem compreender o contexto da pesquisa do pesquisador, bem como as mudanças provocadas na comunidade científica. Além disso, contribuem com o desenvolvimento de políticas públicas, econômicas e científicas (NARIN; OLIVASTRO; STEVENS, 1994).

Moed (2017) classifica os indicadores bibliométricos em quatro categorias; a saber: input, processo, output e impacto. Os indicadores de input estão relacionados à fase inicial do processo de pesquisa e destinam-se à mensuração dos diferentes recursos envolvidos no desenvolvimento das pesquisas. Os indicadores de processo analisam o transcurso, *i.e.*, variáveis e características envolvidas entre o input e o output da pesquisa, caracterizando como se dá o andamento da pesquisa até chegar no produto (output) que pode ser tanto uma produção científica como tecnológica. Baseando-se nos produtos decorrentes da pesquisa, os indicadores de impacto destinam-se a mensurar a influência que o conhecimento produzido tem sobre a sociedade e a comunidade científica.

Considerando o foco de estudo desta tese, na sequência, apresentam-se de forma mais detalhada os indicadores de produção e citação.

Os indicadores de produção refletem a evolução da área por meio dos produtos por ela gerados. Medidos pela contagem de trabalhos e tipologias documentais citadas, são os mais aconselháveis para avaliar o desempenho científico (GABRIEL JÚNIOR, 2014). Os indicadores de produção (ou *output*) são baseados em publicação, como artigo de periódico ou monografia, ou em outros formatos como a produção tecnológica, conjunto de dados de pesquisa, e são direcionados à comunidade científica-acadêmica ou à sociedade e ao público em geral (MOED, 2017).

Salienta-se que o impacto social atende aspectos fora do domínio científico, contemplando aspectos tecnológicos, sociais, econômicos, educacionais e culturais. Os indicadores de impacto tecnológico avaliam o desenvolvimento de novas

tecnologias e estão relacionados ao número de patentes e citações delas provenientes.

Os indicadores sociais, voltados às políticas públicas, são representados por citações em documentos governamentais, organizações nacionais e internacionais, financiamentos, entre outros. Os indicadores econômicos mensuram e analisam o valor e desempenho da produtividade, da economia e a capacidade de inovação. Já os indicadores de impacto educacionais analisam o desempenho com base em visibilidade, palestras, premiações, entre outras. E por sua vez os indicadores culturais avaliam o cenário antropológico e a capacidade de desenvolver no indivíduo a compreensão de suas origens (MOED, 2017; GRÁCIO, 2020).

Muitos estudos de citação adotam as citações de patentes para analisar fluxos de conhecimento de empresa para empresa, ou de outros setores, por exemplo, institutos de pesquisa e academia para empresas (MEYER, 2002; LEYDESDORFF, MEYER, 2003). Guzmán-Sánchez (1999) categorizou os indicadores de patentes em:

- a) **Indicadores de atividade** - analisados sobre três aspectos fundamentais:
 - número e distribuição de patentes solicitadas ou concedidas;
 - produtividade dos inovadores, países, instituições;
 - contagem de citações;
- b) **Indicadores relacionados à primeira geração** – identificam os laços e as interações entre os diferentes campos. Constroem-se a partir dos vínculos estabelecidos pela patente (de inventores ou signatários), as redes de citações, as citações de artigos científicos nas patentes, as cocitações, as colaborações entre inventor, empresa e outras instituições, entre outros;
- c) **Indicadores relacionados à segunda geração** – baseiam-se nas informações presentes no título, no resumo ou no próprio texto. São criados para tratar o conteúdo dos documentos, como as coocorrência de palavras;
- d) **Indicadores relacionados à terceira geração** – são direcionados à inteligência artificial como lógica fuzzy (aprendizado indutivo), algoritmos genéticos e redes neurais artificiais;

- e) **Família de patentes**⁶ – analisada como um indicador independente, visto que a família é, por si só, uma fonte de geração de indicadores e por suas características distintas.

Por sua vez, a Organização de Cooperação e Desenvolvimento Econômico (OCDE)⁷ lançou a segunda edição do Manual de Estatística de Patentes da OECD (PARIS, 2009) que apresenta diretrizes fundamentais para o uso de dados de patentes para mensurar CT&I. Conforme consta no manual, os documentos de patente contêm uma quantidade expressiva de informações para análises estatísticas, esses indicadores são categorizados em:

- a) **Descrição técnica da invenção** - título e resumo, lista de reivindicações, classes técnicas a qual a invenção pertence (classificação internacional de patentes - CIP), arte anterior (determina os limites do que é do domínio público e a que o requerente tem direito em relação às reivindicações), referência de patentes e referências que não são patentes (publicações científicas, anais de conferências, livros e demais);
- b) **desenvolvimento e propriedade da invenção** – inventores e depositantes e seus respectivos endereços;
- c) **histórico do aplicativo** - número da publicação, número do depósito, número da concessão, data de depósito, data de prioridade, data de publicação, lista de designação⁸ (para os pedidos de patentes apresentados usando os procedimentos do Tratado Europeu de Cooperação de Patentes ou da Convenção de Patentes, os solicitantes devem designar os países-membros nos quais a proteção está sendo solicitada), data de recusa ou retirada (indica que a invenção não cumpriu os critérios legais ou que o requerente decidiu suspender o pedido de patente durante o processo de exame), data de

⁶ É um conjunto de patentes depositadas em diferentes países com o propósito de proteger uma mesma invenção, ou seja, são patentes que dividem o mesmo número prioritário (primeira patente da invenção/patente base) (MINISTÉRIO DA ECONOMIA, 2021).

⁷ Criada em 1948, com o nome de Organização Europeia de Cooperação Econômica (OECE), tinha como objetivo recuperar um continente devastado pela Segunda Guerra. Em 1961 passou a chamar de Organização de Cooperação e Desenvolvimento Econômico (OCDE), composta por 36 países que queriam identificar problemas, discuti-los e analisá-los, além de promover políticas para o bem-estar econômico e social. <<http://www.oecd.org/about/>>

⁸ É a lista de países que se pretende proteger uma invenção e que são membros do Tratado de Cooperação em matéria de Patentes (PCT) e da Convenção da União de Paris.

concessão e data de expiração (quando as taxas de anuidade não forem pagas).

Mais recentemente, Jürgens e Herrero-Solana (2017) apresentaram uma categorização mais refinada para os indicadores de patente, visto que compreendem que há uma progressão nas análises patentométricas. Segundo eles, com mais de 95 milhões de patentes publicadas, as informações contidas nesses documentos são importantes para realizar uma vigilância tecnológica de domínios tecnológicos específicos, pois seus dados refletem as atividades de campos de tecnologia com altas desempenho em pesquisa e desenvolvimento. Sendo assim, eles categorizaram os indicadores tecnológicos em:

- a) **Análise de campo único** – como o nome indica, é a análise aplicada a um único campo, normalmente baseada em listas, rankings;
- b) **Análise de campo múltiplo** – também nomeada de análise de referências cruzadas combina distintos campos por meio de matriz. Aplica-se em redes de colaboração. Aqui os indicadores são classificados em:
 - Indicadores de desempenho – tratam da produção de patentes das entidades analisadas e são uteis para monitorar o desempenho tecnológico de empresas, Estados e demais, além de permitir rastrear e elite tecnológica.
 - Indicadores de tecnologia – analisam as classificações das patentes, o que pode inferir os pontos focais tecnológicos de uma organização, a evolução de um setor e os domínios tecnológicos em âmbito micro, meso e macro.
 - Indicadores de valor de patente – fornecem informações sobre o valor econômico das patentes, observando fatores como o tamanho da família de patentes, cobertura geográfica, citações de patentes entre outros.
 - Indicadores de colaboração - indicam os padrões de colaboração das entidades. Gerados com múltiplas análises, podem ser verificados por meio de mapas de redes.

Cativelli (2020) considera que os modelos de categorização dos indicadores de patentes apresentados pelos diversos atores, embora recebam nomenclaturas diversas, permitem realizar a mesma análise dos dados, ou seja, o modelo aplicado não interfere nos resultados. Moed (2017) já afirmava que as patentes (indicador de

produção) também podem ser usadas como indicador de impacto social, visto que traz consigo um valor tecnológico e econômico, indicando o desempenho em C&T do país.

Diante do exposto, compreende-se que as análises baseadas em referências compreendem indicadores de produção e as análises baseadas em citações abrangem indicadores de impacto. Nesse sentido, analisar esses indicadores contribui para conhecer, no cenário tecnológico, o que na literatura científica corresponde à identidade e imagem de citação dos autores.

4.2 IDENTIDADE E IMAGEM DE CITAÇÃO

No processo de construção do conhecimento científico, observa-se que o pesquisador sustenta seu trabalho fundamentando suas ideias em estudos anteriormente desenvolvidos. Esses estudos registrados no trabalho do pesquisador aparecem listados ao final do documento científico, indicando com quais atores⁹ ele dialogou. Sob essa perspectiva, a partir da análise do conjunto de referências elaborado pelo autor, é possível estimar seus principais interesses científicos.

No âmbito das referências, é possível observar que à medida que a produção científica de um pesquisador aumenta, também cresce os atores por ele referenciados, aparecendo, muitas vezes, em várias das suas publicações, o que reflete, segundo Grácio (2020), sua identidade científica de citação.

Na perspectiva de White (2001), a identidade de citação é formada pelo conjunto de autores que um pesquisador cita com recorrência em suas publicações. Segundo ele, os autores citados podem ser divididos em duas classes: os citados em uma única publicação e os que são citados em mais de uma publicação pelo pesquisador. Os citados que reincidem com frequência nas diversas obras de um pesquisador são denominados recitados e formam a sua identidade de citação. Embora White (2001) mencione a possibilidade de que a identidade de citação de um pesquisador possa ser premeditada, em virtude de suas escolhas, considera que os padrões de citação que se formam ao longo da sua trajetória científica acontecem espontaneamente, de forma não planejada.

Ainda sobre a recitação, cabe destacar que ela se desdobra em sincrônica e diacrônica. A primeira pode ser observada quando um pesquisador cita um autor mais

⁹ atores são tratados no estudo como sinônimos de autores e pesquisadores.

de uma vez em apenas um trabalho, já a perspectiva diacrônica é identificada quando observa-se que o citante mencionou o mesmo citado em diversas publicações. A análise de recitação diacrônica pode evidenciar que o autor recitado diversas vezes ao longo da produção científica de um pesquisador certamente tem um elevado grau de importância para esse pesquisador, ou seja, um autor que norteia seu pensamento científico (WHITE, 2001; GRÁCIO, 2020).

Observando que a identidade de citação apresenta apenas uma face da característica científica de um pesquisador, cabe mencionar que ela se completa ao observar o que White (2001) chamou de imagem de citação do autor, estabelecida por quem o cita e com quem é citado conjuntamente, ou seja, é formada pelo conjunto de autores com quem o autor é cocitado.

White (2001) mostrou que o uso das técnicas bibliométricas pode contribuir para identificar essa característica e propiciar uma compreensão mais precisa das relações sociais e cognitivas exteriorizadas nos estudos de redes.

Cronin e Shaw (2002) afirmam que um conjunto de autores que mencionam em seus estudos um pesquisador são denominados de criadores de imagem de citação desses pesquisadores. Isso porque a imagem do citado se estabelece sob a perspectiva de quem cita. Os citantes não planejam de forma consciente a imagem de citação de um pesquisador, ela acontece naturalmente à medida que ele recebe múltiplas citações simultaneamente com outros citados. Sendo assim, a imagem de citação é definida pela comunidade científica.

White (2001) afirma que a imagem da citação muda com o tempo, conforme o trabalho do citado é incorporado por outros autores em novas listas de referências. A imagem de citação, segundo ele, é determinada por outros, exceto na medida em que se contribui para ela por meio de autocitação em listas de referência próprias.

Assim como White (2001), Grácio (2020) também considera que a dinamicidade da ciência contribui não apenas para a mudança da imagem de citação do pesquisador, mas também para sua identidade. Segundo a autora, com o passar dos anos a imagem e a identidade de um pesquisador podem sofrer alterações. A primeira em função das mudanças de interesse dos pesquisadores que os citam e a segunda em virtude da dependência do olhar e da perspicácia da comunidade científica que pode desenvolver outros interesses de estudo e assim buscar novos autores para dialogar com suas ideias.

Além disso, cabe entender que no processo de construção e consolidação da sua produção científica, um pesquisador cita muitos autores e são esses que compõem sua identidade, mas não necessariamente sua imagem, a menos que outros os coloquem lá. Nesse contexto, é válido mencionar que a autocitação, muito usual, afeta a imagem de citação de um pesquisador, visto que ele aparece com os demais autores que compõem sua imagem, aumentando sua contagem de citação. Dentro da comunidade científica, a autocitação, embora comum, recebe muitas críticas por parte dos pesquisadores, visto que ela infla a contagem de citação dos citantes num processo egocêntrico, pois fortalece sua posição dentro da comunidade (WHITE, 2001). No entanto, esse comportamento vem sendo discutido e mudanças estão ocorrendo para dirimir a autocitação excessiva como mencionam Ioannidis e coautores (2019). Segundo Cronin e Shaw (2002) a autocitação é a forma mais dominante de citação, o que reitera a fala de White (2001) de que a autocitação é o “[...] núcleo do núcleo.” da identidade de citação de um indivíduo.

Há quase quatro décadas, Narin e Noma (1985) afirmavam que os dados de citação e referências de patentes em biotecnologia, assim como os artigos de biociência podiam ser usados para mostrar que as técnicas bibliométricas aplicáveis em ambos os domínios eram semelhantes. Assim como as publicações científicas, as patentes também recebem referências e podem ser tanto de documentos não patentes (livros, artigos e demais) como também de outras patentes. Essas referências contribuem para comprovar ou não a novidade em vista do desenvolvimento tecnológico existente.

Assim como para Grácio (2020), analisar o conjunto de referências de determinada obra permite identificar o interesse científico de determinado autor, observa-se que tal afirmação também se aplica aos documentos de patentes, visto que, tanto os inventores como os examinadores fazem muitas citações a documentos de patentes e não patente também. Moura (2009) aponta que, embora na literatura científica, os avaliadores e editores têm liberdade para sugerir ou excluir citações, no documento de patente, é o examinador que tem um papel mais proativo, pela inclusão e exclusão de referências, o que torna o processo de citação muito mais complexo. Além disso, Bassecouard e Zitt (2004) consideram que o ato de citar em literatura científica é mais seletivo e restrito, já em documentos de patente, busca-se a exaustividade para garantir que tal pedido atende os critérios de patenteabilidade.

Salienta-se, ainda, que a análise de citação em patentes também defronta as dificuldades de se conhecer os motivos que levam um pesquisador a citar os trabalhos de outros, pois as citações em patentes podem refletir as diferentes motivações de dois indivíduos que são: os inventores e os examinadores de patentes (MEYER, 2000).

Considerando que os motivos para citar sejam diversos, cabe recordar o que pondera White (2001): o elemento fundamental para o ato de citar está relacionado ao reconhecimento e a relevância científica e tecnológica, em se tratando de patentes. Nesse contexto, sob a perspectiva de White (2001) e de Cronin e Shaw (2002), entende-se que, no cenário das patentes, a imagem de citação se dá no conjunto de todos os autores e demais patentes que esta patente foi citada simultaneamente na lista de referências. Sendo a imagem de citação determinada pela ótica da comunidade de inventores e de examinadores. Por sua vez, a identidade de citação das patentes também acontece sob o olhar dos inventores e examinadores que citam os documentos de patente e não patente a fim de comprovar a novidade ou refutar.

4.3 ANÁLISES RELACIONAIS DE CITAÇÃO

Os estudos de citação, muito voltados ao impacto e a visibilidade dos autores, têm se mostrado um método muito viável para avaliar o desempenho científico. Partindo do pressuposto que as citações podem ser compreendidas como representação de uma comunidade discursiva ou de uma proximidade teórica-metodológica entre pesquisadores, a análise desse indicador explicita as relações existentes entre eles em um determinado domínio. Inferindo a ideia de que a citação carrega consigo essa proximidade teórica entre os autores, Grácio (2020) elucida essas relações apontando para o seu desdobramento em dois tipos de análise: univariáveis e relacionais. Segundo a autora, em análises univariáveis cada sujeito do estudo é investigado individualmente, conforme o indicador determinado. Por outro lado, nas análises relacionais procura-se reconhecer as relações de proximidade teórico-metodológica entre os pesquisadores (WHITE, 2001; ROSTAING, 1996).

Dessa forma, compreende-se que as análises de citação univariadas não proporcionam a identificação da conectividade entre os citados, sendo necessário dados relacionais que promovem uma análise simultânea entre duas referências, indicando proximidade ou distanciamento entre autores ou documentos. Sendo assim,

as análises relacionais têm como propósito identificar a “[...] co-ocorrência de um par de referências [...]” para aferir com que frequência eles aparecem citados juntos (GRÁCIO, 2020, p.73).

Os estudos relacionais de citação foram categorizados por Marshakova (1981) em cocitação (prospectivo) e acoplamento bibliográfico (retrospectivo). A cocitação, a partir da análise de frequência com que dois autores/documentos são citados juntos, permite inferir a influência na área e a proximidade temática. Por sua vez, o acoplamento bibliográfico (KESSLER, 1965) mede a ligação entre dois documentos com base no número de documentos que eles têm citados em comum. Desse modo, compreende-se que o acoplamento bibliográfico se concentra em grupos de artigos que citam documentos em comum, enquanto a cocitação se concentra em referências que vêm frequentemente em pares.

É importante salientar que embora em documentos de patentes as citações podem ser decorrentes de dois atores diferentes (inventores e examinadores), nos estudos de citação nesse universo não há distinção entre citação de um e de outro ator, mas sim entre tipo de documento citado: patente e não patente. Isso pode acarretar duas análises utilizando o mesmo método, ou seja, uma análise de acoplamento bibliográfico para citações às patentes e outra para citação a documentos não patente.

4.3.1 Acoplamento bibliográfico

O termo acoplamento bibliográfico (AB) foi postulado por Kessler (1963), que afirmou que vários artigos científicos têm uma relação significativa entre si quando apresentam uma ou mais referências em comum. Em suas indagações, o autor expressou que a bibliografia (de artigos) é uma maneira pela qual o autor pode indicar o ambiente intelectual dentro do qual ele opera, e se dois artigos apresentam bibliografias semelhantes há uma relação implícita entre eles. Desse modo, dois documentos são considerados acoplados bibliograficamente quando neles são identificados um ou mais itens em comum na lista de referência (EGGUE; ROUSSEAU, 2002).

A fim de elucidar melhor o método apresentado por Kessler (1963), o acoplamento bibliográfico é definido da seguinte forma: considere dois artigos A e B e suas listas de referências. Essas listas de referência consistem em conjuntos de

publicações (documentos citados) e o acoplamento bibliográfico acontece quando há similaridade de citação entre os documentos A e B (Figura 5).

Figura 5 – Artigos A e B acoplados bibliograficamente pelos documentos 3, 4, 5 e 6



Fonte: elaborado pela autora.

Com base na Figura 5 podemos afirmar que o artigo A e o artigo B estão acoplados bibliograficamente porque ambos citam os documentos 3, 4, 5 e 6. Cabe mencionar que a força de acoplamento bibliográfico que acontece entre os artigos A e B não se altera com o passar dos anos, sendo considerada por Small (1973) como uma relação fixa e permanente, visto que depende das referências contidas nos documentos acoplados.

Em suas obras, Kessler (1963; 1965) deixa claro que o conjunto de referências listadas pelos pesquisadores em suas publicações, notabiliza o ambiente intelectual que eles estão inseridos e, dessa forma, se duas publicações desses pesquisadores apresentam referências análogas, significa que pode haver uma similitude intelectual entre eles.

Dessa forma, como mencionam Lucas, Garcia-Zorito e Sanz-Casado (2013), pode-se dizer que a força de acoplamento de dois artigos depende da quantidade de referências comuns entre eles. Sendo assim, é possível dizer que a força de acoplamento entre os artigos A e B (Figura 5) é igual a quatro, visto que eles estão acoplados bibliograficamente por quatro documentos.

Castanha, Bufrem e Bochi (2020) lembram que o acoplamento bibliográfico pode ser aplicado de acordo com a metodologia adotada pelo pesquisador, promovendo outras possibilidades de força de acoplamento conforme o domínio. Essa afirmação fica evidente quando buscamos, retrospectivamente, os estudos de Zhao e

Strotmann (2008; 2014), pioneiros na proposição desse método para o âmbito (nível de agregação) dos autores, a fim de mapear a atividade científica entre eles e, com isso, contribuir para uma visibilidade mais factual das redes baseadas em autores.

Com base no exposto, compreende-se que o método de acoplamento bibliográfico se aplica perfeitamente em estudos voltados à tecnologia e à inovação. Essa constatação é corroborada pelos seguintes estudos: Huang, Chiang e Chen (2003), Yeh *et al.* (2012), Ávila-Robinson e Sengoku (2017) e Kuana, Chenb e Huang (2019).

A análise na tese adota a similaridade entre as patentes a partir das patentes anteriores citadas em comum em lugar dos inventores ou depositantes. Dessa forma, utiliza-se o AB como proposto por Kessler (1963, 1965) e não na perspectiva (ABA) de Zhao e Strotmann (2008, 2014). O nível de mensuração aqui é de documentos e não entre autores (inventores quando traduzido por analogia para o âmbito das patentes).

Huang, Chiang e Chen (2003) já consideravam que estudos aplicados às patentes têm como intenção encontrar uma conexão entre tecnologias, países e empresas. Dessa forma, o autor considera que se as patentes A e B compartilham determinadas patentes em suas citações, elas estão bibliograficamente acopladas e isso, do ponto de vista da indústria pode indicar uma aliança de cooperação ou concorrência.

Yeh e coautores (2012) consideram que o método proposto por Kessler (1963) foi adequado para avaliar a relevância da citação de patente. No estudo, os autores usaram a abordagem de acoplamento bibliográfico (BC) para filtrar as citações de patentes denominadas por eles de irrelevantes. Lo (2008) aplicou o acoplamento bibliográfico para revelar as possíveis ligações tecnológicas existentes entre as principais organizações de pesquisa em engenharia genética. Essas organizações foram definidas pela produtividade e identificadas pela contagem de patentes e pela Lei de Bradford. Para um período de 12 anos, o autor analisou as patentes citadas nas patentes concedidas às 40 principais organizações pelo United States Patent and Trademark Office (USPTO), para estabelecer a proximidade entre elas. Dos 780 pares de acoplamento formados pelas 40 organizações, o estudo mostrou que 151 (19,4%) tinham menos de 10 patentes compartilhadas, 261 (33,5%) pares compartilharam de 10 a 29 patentes citadas e apenas 19 (2,4%) pares apresentaram maior força de acoplamento, ao compartilhar mais de 200 patentes citadas. O estudo permitiu

identificar que entre os pares de empresas com maior força de acoplamento existe uma proximidade teórica sobre tecnologia de sequenciamento genético e que essas organizações são sediadas na França, indicando um forte interesse e possível potencial tecnológico nessa temática e região.

Conforme podemos ver no estudo de Grácio (2020), o acoplamento bibliográfico permite identificar o cerne da pesquisa em determinado domínio, além dos seus pesquisadores e as publicações mais importantes. Essa contribuição também fica evidente quando se incorpora esse método para análise das patentes.

Figura 6 – Ilustração do acoplamento bibliográfico entre patentes por meio do compartilhamento de referências entre elas

Referências Patentes Acopladas	Referências								
	Ref 1	Ref 2	Ref 3	Ref 4	Ref 5	Ref 6	Ref 7	Ref 8	Ref 9
Patente 1		X	X		X		X	X	
Patente 2	X	X		X	X	X		X	X

Fonte: adaptado de Grácio (2016).

Na Figura 6 as patentes 1 e 2 são acopladas pelas referências Ref 2, Ref 5 e Ref 8, o que indica que a frequência de acoplamento bibliográfico entre essas duas patentes é igual a 3.

4.3.2 Cocitação

Cunhado por Small (1973), a análise de cocitação tem como proposta analisar a conectividade entre dois documentos, com base na frequência com que eles são citados juntos por outros documentos.

Diferentemente do acoplamento bibliográfico que liga documentos/autores a partir de duas citações em comum, a cocitação revela a ligação de dois ou mais documentos citados conjuntamente em uma lista de referência das publicações de autores citantes. Desse modo, para que dois documentos sejam fortemente cocitados, é importante que um número expressivo de autores cite os dois trabalhos simultaneamente (Figura 7) (SMALL, 1973; GRÁCIO, 2016).

Figura 7 – Documentos 3 e 4 cocitados em função dos artigos citantes A, B e C

Artigo A cita	Artigo B cita	Artigo C cita
1	3	3
2	4	4
3	5	22
4	6	60
7	10	56
8	11	87
9	12	44

Fonte: elaborado pela autora.

Conforme vemos na Figura 7, os documentos 3 e 4 são cocitados porque ambos são referenciados nos artigos A, B e C. A ilustração evidencia que a interlocução e a proximidade teórica e/ou metodológica entre os documentos 3 e 4 não foi estabelecida pelos seus próprios autores, mas pelos autores que os citaram (artigos A, B e C), elucidando a ideia de Small (1973) de que a partir da força de cocitação é possível prever o grau de associação entre os pares de documentos citados. Além disso, observa-se que a força de ligação entre os pares de documentos pode crescer com o tempo, a medida em que eles aparecem cocitados em documentos posteriores (MARSHAKOVA, 1981).

Marshakova (1981) apontou que essa análise prospectiva também pode ser aplicada para autores, possibilitando analisar e identificar possíveis proximidades temáticas ou metodológicas entre eles. Idealizadores da análise de cocitação de autores, White e Griffith (1981) conceituaram a autoria como o conjunto das obras de um autor.

Os autores, assim como Grácio (2020), afirmam que a análise de cocitação de autores permite identificar a proximidade ou o distanciamento relacional entre pares deles por meio das redes. Pares de autores citados repetidas vezes em documentos subsequentes indicam maior força de cocitação, ao passo que pares de autores citados com pouca frequência aparecem distantes nas redes, indicando menor força de cocitação.

Para analisar a proximidade intelectual entre autores, McCain (1990) estabeleceu seis passos importantes para o desenvolvimento da análise de cocitação de autores; que são:

1. seleção dos autores;
2. recuperação da frequência de cocitação;
3. compilação da matriz de cocitação com valor absoluto;
4. conversão em matriz de correlação (normalizada);
5. análise multivariada da matriz de correlação (normalizada);
6. interpretação e validação dos dados.

McCain (1990) considera importante, ao mencionar o sexto passo, consultar os especialistas das áreas analisadas, a fim de caracterizar os princípios epistemológicos que validam o estudo e assim contribuir para melhor compreender o domínio estudado.

A partir do que foi apresentado pelos autores que contribuem para a fundamentação teórica sobre cocitação, observa-se que estudos aplicados ao cenário tecnológico têm sido desenvolvidos com sucesso, como podemos constatar pelas pesquisas: Mogee e Kolar (1998), Breitzman e Mogee (2002), Lai e Wu (2005), Schultz e Joutz (2010), Erdi *et al.* (2013), Meyer *et al.* (2014), Rodriguez *et al.* (2015), entre outros.

Breitzman e Mogee (2002) sinalizam a importância de analisar as patentes dentro de cenário organizacional. Segundo os autores, ter uma boa gestão de propriedade industrial permite identificar o real valor do portfólio de patentes de uma empresa, saber que tecnologias dispenderam muitos recursos e apresentaram um lucro pouco expressivo, identificar que outras empresas e países apresentam um potencial tecnológico sob a mesma tecnologia, entre outras observações. Para eles, muitas empresas, principalmente as que apresentam um portfólio muito grande, podem não ter um controle exato sobre o número de patentes que possui, em quais tecnologias investiu e em que países têm proteção.

Nesse contexto, Breitzman e Mogee (2002) consideram que a cocitação é um dos métodos para obter melhor controle de gestão das patentes. Entendem que existe uma ligação de cocitação entre duas patentes se ambas forem citadas em uma patente posterior e que uma grande quantidade de cocitação entre elas é evidência

de que estão estreitamente relacionadas. Moguee e Kolar (1998) também tinham apontado essa característica e afirmado que documentos cocitados podem se juntar a outros documentos com os quais também são cocitados e, assim, formar um cluster, conectados uns aos outros. Segundo eles, os clusters que se formam a partir da análise de cocitação de patentes correspondem, geralmente, a principal área de competência tecnológica daquele domínio.

Ainda sobre essa abordagem, apontando como uma área emergente com potencial para gerar novos produtos ou aprimorar os já existentes, Schultz e Joutz (2010) aplicaram a análise de cocitação às patentes do campo da nanotecnologia, a partir da qual definiram agrupamentos de patentes relacionadas. Trabalharam com as 101 patentes mais citadas no campo e desenvolveram uma rede de cocitação em que identificaram 7 agrupamentos sobre variados temas relacionados à nanotecnologia, cobrindo desde dispositivos de armazenamento e recuperação de informações a energia radiante, prospectando um elevado potencial tecnológico na computação.

Esses estudos reiteram o que Marshakova (1981) e Hjørland (2002) apontam sobre a eficácia da aplicabilidade da cocitação em diferentes âmbitos (documentos, autores, periódicos, países), proporcionando conexões mais reais e detalhadas dos elementos dessas dimensões. Grácio (2020) considera ainda que a análise de cocitação revela as similitudes existentes entre as ideias dos citados, mas pode também revelar as dicotomias de pensamentos entre eles.

Embora o método seja aplicado sob qualquer perspectiva, o mais usual é a análise de cocitação de autores, empregada para evidenciar a estrutura intelectual de um campo científico. Essa abordagem foi proposta por White e Griffith (1981) e busca revelar a frequência com que um autor está ligado a outro nas publicações da comunidade científica, não diferenciando quais os trabalhos desses autores foram citados conjuntamente. Esse método, como apresentado no estudo de Bayer, Smart e McLaughlin (1990), tem por pressuposto a premissa que quanto maior a frequência em que dois autores são cocitados, maior são os laços teórico-metodológicos entre eles.

Elucubrando os teóricos na temática de análise de cocitação de autores, Grácio (2020) afirma que tal método promove uma melhor visualização da estrutura intelectual de uma comunidade científica. Esse método pode ser aplicado à propriedade industrial, a fim de identificar essa estrutura intelectual no cenário tecnológico. Contudo, observa-se que a aplicação do estudo no ambiente tecnológico

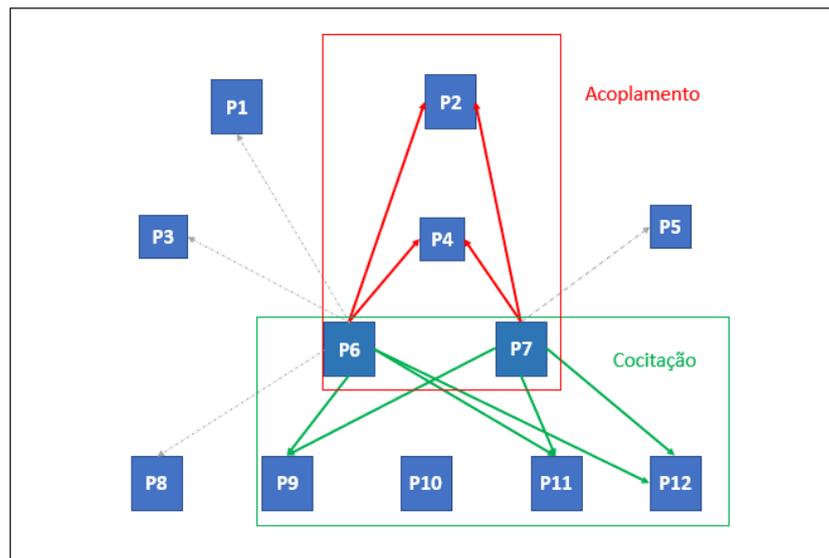
tem explorado outros campos de informação dos documentos de patente que não os inventores. Isso certamente pode estar relacionado ao forte valor monetário que está atrelado ao documento, provocando maior interesse em analisar aspectos como as tecnologias emergentes, os países com maior potencial tecnológico, entre outros.

Nesse contexto, observa-se que as patentes estão cada vez mais ganhando notoriedade nos estudos das mais diversas áreas do conhecimento, tendo como método de análise os estudos métricos da informação. Os primeiros estudos nesse instrumento aplicando as metrias pertencem a autores com Narin (1994), Glänzel e Meyer (2003), Leydesdorff (2008a, 2008b), entre outros estudiosos. Esses estudos pioneiros de métricas aplicadas às patentes se desenvolveram consistentemente, estimulando a aplicação de outras metodologias, como visto nos estudos de Mogee e Kolar (1998), Breitzman e Mogee (2002), Lai e Wu (2005), Schultz e Joutz (2010).

A fim de recuperar referencial teórico abrangente, incluindo literatura clássica, mas também contemporânea para fundamentar a presente pesquisa, realizaram-se buscas nas bases de dados Web of Science (WofS) e Scopus, utilizando os termos “bibliographic coupling” and “co-citation” and “patent*” em todos os campos, trabalhos que abordassem a metodologia aplicada nesse estudo. A escolha do uso da expressão booleana “and” teve como intenção recuperar as pesquisas que aplicam simultaneamente ABA e ACA no seu universo, a fim de identificar quantitativamente os estudos que aplicam o mesmo método.

Foram recuperados nove trabalhos na Web of Science (WoS) e cinco na Scopus, sendo que os cinco trabalhos localizados nessa base estão contemplados também na WoS. Desses nove trabalho recuperados, três se mostraram pertinente para essa pesquisa, sendo eles: Huang, Chen e Dong (2011), que aplica a metodologia na perspectiva da energia fotovoltaica com o objetivo de analisar o que eles chamam de “citações faltantes”; Ávila-Robinson e Sengoku (2017), que aplica as metodologias sob o domínio das células-tronco pluripotentes induzidas (iPS), relacionando os domínios ciência (literatura científica) e tecnologia (patentes); Van Raan (2017), com a revisão do estado da arte dos estudos métricos em patente, aplicando, na sequência, a análise de acoplamento bibliográfico e de cocitação como nova técnica, a fim de visualizar as relações citante e citado nas referências científicas não-patentes (citadas nas patentes).

Figura 8 – Ilustração da rede de acoplamento bibliográfico e de cocitação entre patentes



Fonte: elaborado pela autora.

Em diagrama, a Figura 8 ilustra, simultaneamente, as conexões que se estabelecem para o desenvolvimento tecnológico, pela aplicação das análises de acoplamento bibliográfico e de cocitação, as quais permitem identificar e explicitar as relações entre patentes. As patentes P6 e P7 estão bibliograficamente acopladas pelas patentes P2 e P4 (indicação em vermelho). Por outro lado, foram cocitadas nas patentes P9, P11 e P12 (em verde).

Grácio (2020) considera que as duas análises relacionais são muito aplicadas para identificar e compreender a comunicação científica, as frentes de pesquisa, bem como as relações intelectuais dos diversos domínios tecnológico-científicos.

5 METODOLOGIA

Este capítulo apresenta a descrição do percurso metodológico realizado no desenvolvimento desta tese, contemplando a abordagem, tipo de pesquisa, os processos de busca, recuperação e tratamento dos dados, a fim de alcançar os objetivos propostos nesta pesquisa: compreensão da estrutura tecnológica e científica que fundamenta a produção das patentes em células-tronco.

5.1 ABORDAGEM E TIPO DA PESQUISA

A pesquisa é baseada na abordagem quantitativa, visto que são utilizados os métodos estatístico e métrico, mais especificamente a patentometria e a bibliometria, especialmente pela aplicação dos procedimentos relativos à cocitação e ao acoplamento bibliográfico em documentos de patente, além da contextualização com base na literatura da área. Esse tipo de pesquisa fornece uma descrição mais precisa e objetiva dos dados do objeto estudado.

O tipo da pesquisa é descritivo, entendido como o estudo que tem como característica principal a definição de objetivos e informações sobre o tema estudado, propiciando que se conheça os fenômenos da pesquisa (MARCONI; LAKATOS, 2009; GIL, 2010). Tem como princípio a análise detalhada das características do objeto estudado (HERNÁNDEZ SAMPIERI; FERNÁNDEZ COLLADO; BAPTISTA LUCIO, 2013).

5.2 DELINEAMENTO DA PESQUISA

Será empregada a patentometria, que tem como principal atributo a análise do uso e das características dos documentos de patentes, associada a procedimentos consolidados da bibliometria, mais especialmente estudos de cocitação e acoplamento bibliográfico. Considera-se que os estudos patentométricos analisam os documentos de patentes a fim de conhecer as atividades tecnológicas e inovadoras dos países, áreas de conhecimentos e instituições, permitindo, assim, conhecer o grau de inovação das empresas e países, entender como o conhecimento científico se transforma em conhecimento tecnológico e observar a capacidade competitiva das indústrias (MARTÍNEZ, 2011).

Os indicadores patentométricos demonstram com maior precisão as atividades de inovação e tecnologia de um país, de instituições públicas ou privadas, sejam de ensino ou grandes empresas. Proporcionam compreender as relações entre universidades e empresas no que tange as atividades de CT&I, além de permitir aferir o grau de investimentos nessas atividades, ao longo dos anos (GUZMÁN-SÁNCHEZ, 1999; MACIAS-CHAPULA, 1998).

Como menciona Hjørland (2002), é na combinação das abordagens de análise de domínio (AD), que se fortalece a identidade da área e enriquece as relações entre teoria e prática. A fim de propor uma abordagem tecnológica a partir deste estudo, são utilizadas as abordagens 5 e 8 de AD, a saber: produção e interpretação de estudos bibliométricos e estudos epistemológicos e críticos de diferentes paradigmas, pressupostos e interesses em domínios.

5.3 ESCOLHA DA BASE DE DADOS E ESTRATÉGIA DE BUSCA

Para o desenvolvimento exitoso da pesquisa, a escolha da base de dados e a definição das expressões de busca são etapas fundamentais da construção metodológica. Nesta seção, são descritas a fonte de coleta e as estratégias utilizadas.

5.3.1 Fonte de Coleta dos Dados

Utiliza-se a base de patentes Derwent Innovation Index (DII) como fonte para a coleta dos dados, em função de esta indexar documentos de patentes e suas citações em âmbito mundial, possibilitando, assim, uma cobertura mais ampla e representativa da produção tecnológica sobre o tema células-tronco.

Disponibilizada no Portal de Periódicos da CAPES, essa base de patentes é assiduamente utilizada em estudos de indicadores de CT&I, uma vez que é considerada uma das mais importantes fontes de indexação de informações provenientes de patentes (MARICATO, 2010; CONSONI, 2017; MAGNUS, 2018). Fornece acesso a mais de 30 milhões de invenções descritas em mais de 65 milhões de documentos de patentes, atendendo mais de 48 autoridades emissoras de patentes (escritórios de patentes), desde 1963. Cabe ressaltar que a versão utilizada foi a DII, disponibilizada pelo Portal de Periódicos da CAPES.

5.3.2 Procedimento para Coleta dos Dados

A expressão de busca utilizada foi “*Stem Cell*” a fim de recuperar todos os documentos que apresentam o termo, sem limitação pelo tipo de célula-tronco. Esclarece-se que esta expressão foi considerada a mais adequada, após a realização de pré-testes que buscaram identificar a melhor estratégia de recuperação do universo a ser analisado (conjunto de patentes concedidas sobre célula-tronco), a fim de garantir a representatividade dos resultados obtidos pela pesquisa. A escolha da busca utilizando o termo células-tronco em inglês, decorre do fato de ser o idioma mais abrangente em âmbito tecnológico e científico mundial. Desse modo, conjectura-se que permite a recuperação com maior cobertura do universo analisado.

Para a realização desse pré-teste, utilizou-se a pesquisa avançada, rótulo do campo TS, que compreende a pesquisa por termos de tópicos nos campos Título e Resumo nos registros de patentes. Com o intuito de contemplar o maior número de documentos de patentes, utilizou-se as “aspas” e o código coringa (*), recuperando um total de 42.802 documentos que abordam a temática. A expressão de busca ficou TS=(“Stem Cell*”).¹⁰

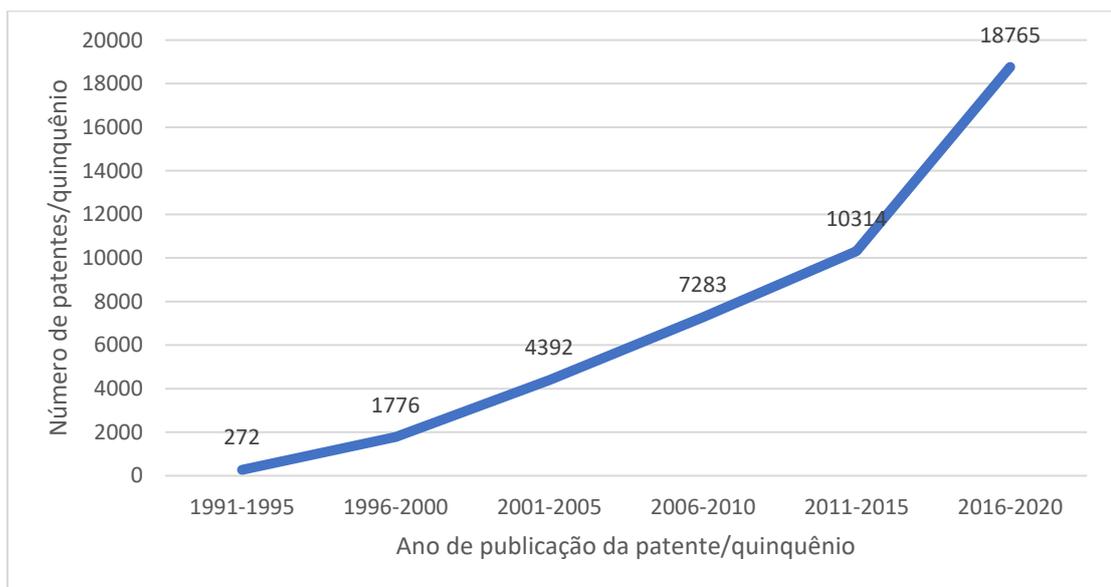
Para esta pesquisa foi delimitado um intervalo temporal de 5 anos, compreendendo dados do período de 2016 a 2020, correspondendo a um total de 18.765 documentos. Price (1969) propôs a lei do elitismo, expressão matemática destinada a identificar o tamanho da elite de pesquisa (os autores mais profícuos) de um campo do conhecimento, formulada em função da quantidade (total) de autores atuantes no campo analisado e expressada como a raiz quadrada ($\sqrt{\quad}$) do número total de autores. A fim de obter para esta pesquisa uma amostra representativa da elite tecnológica do universo de patentes analisado, aplicou-se a lei do elitismo de Price para o total de 18.765 documentos de patentes, ou seja, calculou-se a raiz quadrada dessa quantidade de documentos, resultando em 137 patentes. A seguir, ordenaram-se os 18765 documentos de patentes do mais citado para o menos citado, a fim de identificar as 137 mais citadas. Contudo, o 137º documento foi citado 19 vezes, mesma quantidade (19) de citações das 7 patentes seguintes. Desse modo, o universo de estudo é formado por 144 documentos (Apêndice B).

¹⁰ Usa-se caracteres coringas como (*) para encontrar formas no plural e flexionadas de palavras. Para pesquisar uma expressão exata, use aspas (“”).

Salienta-se que não foi limitado o espaço geográfico, pois os dados são analisados em âmbito mundial. Desse modo, o corpus analisado é composto pelas 144 patentes, citadas ao menos 19 vezes cada e suas respectivas referências às patentes citadas (4915 patentes citadas).

A janela temporal de cinco anos (2016-2020) se justifica por ser o período que apresentou um corpus consistente e representativo da estrutura tecnológica contemporânea no campo das células tronco, adequado para o desenvolvimento da pesquisa, como se pode verificar no gráfico 1. Ademais, considerando que um domínio é dinâmico e está sempre em evolução no tempo e no espaço (HJØRLAND, 2017), é imprescindível que se estabeleça uma janela temporal para caracterizá-lo de forma mais significativa e precisa.

Gráfico 1- Distribuição de frequência em quinquênios das patentes publicadas na base DII no período de 30 anos.



Fonte: elaborado pela autora.

Conforme observa-se no Gráfico 1, o número de pedidos de patentes é 69 vezes maior no último quinquênio (2016-2020) em relação ao primeiro (1991-1995), indicando que há um potencial tecnológico no que tange o tema células-tronco.

Conforme os critérios mencionados acima, levando em consideração as famílias de patente, nesta pesquisa, utilizou-se a primeira patente que os metadados apresentaram e não a patente original. No entanto, considera-se importante listar os

escritórios que mais apresentaram pedidos de patente sobre a temática células-tronco.

O Tabela 1 indica os escritórios onde foram realizados os pedidos das patentes originais. Observa-se que os escritórios dos Estados Unidos e Canadá correspondem mais de 60% dos países de pedido de proteção, indicando seu potencial tecnológico no assunto. É possível que esse fato se dê pelo fato de o Projeto Genoma Humano (PGH) ser um consórcio internacional de pesquisas liderado pelos Estados Unidos e que tem como demais colaboradores países como o Canadá, além da França, Alemanha, Japão e Inglaterra.

Tabela 01 – Distribuição de quantidade e percentual dos escritórios do depósito do pedido original das patentes

Escritórios (Código ISSO 3166)	Quant. patente	%
Canadá (CA)	47	32,64
Estados Unidos (US)	41	28,47
China (CN)	20	13,89
Escritório Europeu (EPO)	17	11,81
Austrália (AU)	6	4,17
Grã-Bretanha (GB)	3	2,08
República da Coreia (KR)	3	2,08
Taiwan (TW)	2	1,39
Índia (IN)	2	1,39
Singapura (SG)	1	0,69
Israel (IL)	1	0,69
Japão (JP)	1	0,69
Total	144	100,00

Fonte: elaborado pela autora.

Considerando que a citação à patente se refere à família de patente e não a uma única patente, optou-se por analisar o primeiro número da patente que os metadados apresentaram. Por isso, ao longo da pesquisa veremos a presença de patentes WO, que correspondem aos pedidos feitos via Tratado de Cooperação de Patentes (PCT).

5.3.3 Tratamento e Análise dos Dados

Após a recuperação do conjunto de patentes a ser analisado na base DII, os dados foram exportados em formato csv (uma das opções da base) no dia 29 de

agosto de 2021. Contudo, observou-se que algumas citações haviam sido suprimidas devido ao espaço das células no software Excel, o que levou a uma nova coleta. Para que nenhuma informação se perdesse, percebendo que o número de citações que as patentes recebem se altera mês a mês, realizou-se a exportação dos dados dia 20 de janeiro de 2022 e, nesse dia, criou-se uma lista de conteúdo das 144 patentes utilizadas na pesquisa. A base passou por atualizações e algumas melhorias no processo de exportação dos dados foram observadas.

As referências de documentos não-patente foram identificadas no campo referências citadas (CR), listadas pelo examinador e pelo inventor. Salienta-se que parte das mais de 900 referências de uma patente foi suprimida quando a primeira exportação foi realizada em csv. Dessa forma, optou-se por fazer uma exportação em formato txt. personalizada, ou seja, apenas com os campos necessários à pesquisa, como PN – número da patente prioritária¹¹, CR – referências citadas não-patentes e CP – citações às patentes.¹² Com o intuito de tornar as descrições anteriormente mencionadas mais visíveis, criou-se o Quadro 1, relacionando os objetivos do estudo com os indicadores que serão utilizados nesta pesquisa.

Quadro 1 – Relação dos objetivos específicos e indicadores do estudo.

#	OBJETIVOS ESPECÍFICOS	INDICADORES
a	Identificar o referencial tecnológico presente no universo das patentes pesquisadas;	Citação a patente (CP)
b	Caracterizar o referencial científico (citações à literatura não-patente), adotado no conjunto de patentes de células-tronco da pesquisa;	Citação à literatura não patente (CR)
c	Identificar os agrupamentos da produção tecnológica relacionadas a células-tronco, caracterizados pelas proximidades tecnológicas e científicas das patentes indexadas na base DII, evidenciados pelas similaridades nas suas referências e mensuradas pelo acoplamento bibliográfico;	Acoplamento Bibliográfico
d	Descrever os agrupamentos tecnológicos e científicos identificados a partir das imagens (citações simultâneas) das patentes em células-tronco.	Cocitação
e	Reconhecer os métodos bibliométricos de acoplamento bibliográfico e cocitação entre patentes para identificação de domínios científicos-tecnológicos.	Análise Epistêmica e Crítica

Fonte: elaborado pela autora.

¹¹ Conceitos apresentados na nota 4.

¹² As tags em inglês são: CP - Cited Patent(s) e CR - Cited Article(s). Ver: https://images.webofknowledge.com/WOKRS535R111/help/DII/hs_dii_fieldtags_patents.html

5.3.3.1 Procedimentos para análise de cocitação de patentes citadas

Para o processo de análise de cocitação utilizou-se o software PHP¹³ a fim de padronizar as famílias de patentes, pois observou-se que a não padronização poderia causar um resultado enviesado. Assim, as famílias de patentes foram substituídas pela primeira patente composta no campo PN (Figura 9).

Figura 9 – Padronização das famílias de patentes localizadas no campo PN

```

FN Clarivate Analytics Web of Science
VR 1.0
PT P
FN WO2019051380-A1; TW201919673-A; AR113011-A1; AU2018330322-A1; CA3075268-A1; KR2020053531-A; CN111148531-A;
FN Clarivate Analytics Web of Science
VR 1.0
PT P
FN WO2019051380-A1; WO2019051380-A; WO2019051380-A1; WO2019051380-A1; WO2019051380-A1; WO2019051380-A; WO2019051380-A;

```

Fonte: Elaborado pela autora.

Nos estudos que adotam, como procedimento metodológico, a análise de cocitação de autores, entre eles Grácio e Oliveira (2013; 2014) e Grácio (2020), é usual observarmos a menção à utilização de uma matriz quadrada e simétrica com a frequência de cocitação entre os autores mais citados. Contudo, por ser inviável a adoção desse procedimento para esta pesquisa, uma vez que há mais de 4900 patentes citadas, gerou-se uma matriz de dimensão de 144 linhas e 4915 colunas (Figura 10) para a análise de cocitação, em que as patentes localizadas no campo PN (144 citantes) foram colocadas em linha e as patentes citadas, localizadas no campo CP (4915 patentes citadas) foram realocadas em coluna.

Figura 10 – Recorte do processo de transformação dos dados em matriz

PN	WO2019051380-A1;	WO2019051380-A;
CP	WO2019051380-A1	WO2019051380; EP2494865; WO2003051379; WO2014152484;
	EP2494865-A2	
	WO2003051379-A1	
	WO2014152484-A1	

Fonte: elaborado pela autora.

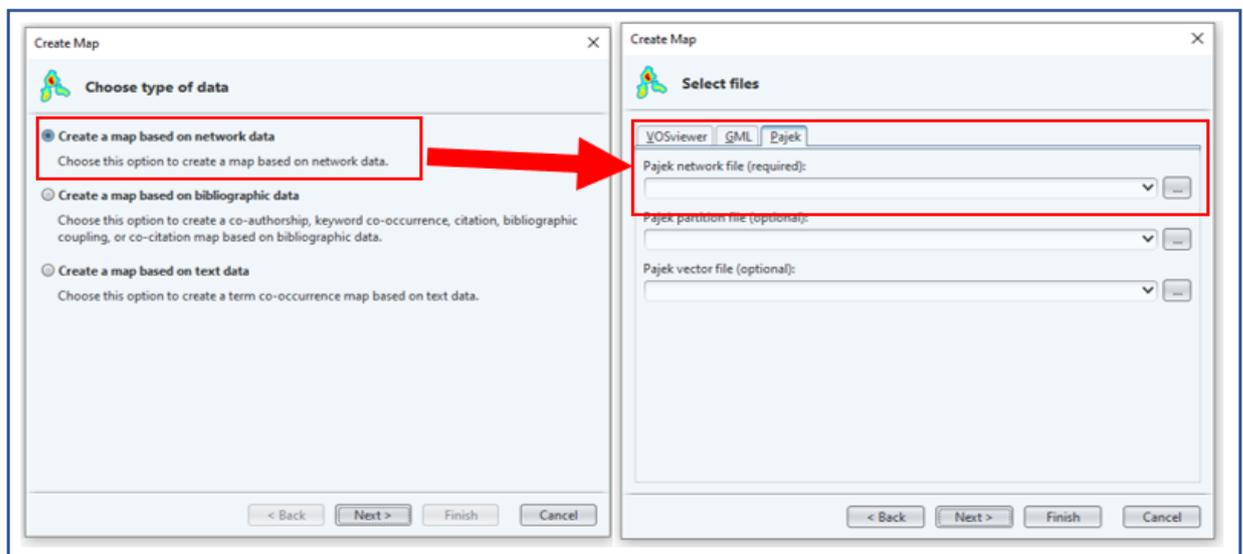
Para se viabilizar a construção da rede de cocitação entre as referências de patentes citadas, os arquivos que estavam em txt. foram convertidos em arquivo Pajek

¹³ Os scripts do PHP podem ser consultados por meio do link:
https://drive.google.com/drive/folders/1IWLBXBkl-LkqENBoYM_KUy7ObyR5MoXI?usp=sharing

(.net), por meio da ferramenta de conversão da Base de Dados Referenciais de Artigos de Periódicos em Ciência da Informação (Brapci), para que pudessem ser lidos no software VOSviewer (CWTS, 2021). A Ferramenta da Brapci realiza essa conversão, identificando as colunas como item citante e as referências citadas de cada item por linha no excel.

Já o software VOSviewer (CWTS, 2021) é uma ferramenta que permite elaborar e visualizar, por meio de campos específicos, redes bibliométricas de citação, cocitação, acoplamento bibliográfico, entre outras análises. A ferramenta acessa os registros das referências para as análises de cocitação e de acoplamento bibliográfico de autores, a partir dos dados exportados de bases como Scopus e Web of Science. Ademais, constrói redes bibliométricas a partir de arquivos do tipo VOSviewer, GML e Pajek. Nesta pesquisa, para que os arquivos .net, oriundos da conversão realizada anteriormente, pudessem ser lidos corretamente, utilizou-se a opção “Criar mapa baseado em dados de redes” e, na sequência, escolheu-se a aba Pajek como ilustrado na Figura 11.

Figura 11 – Identificação dos campos utilizados para leitura dos dados da rede de cocitação



Fonte: elaborado pela autora.

Após esse procedimento, o software identificou uma rede com 4915 patentes citadas. Porém, a fim de obter uma rede mais elucidativa para a análise de cocitação entre as patentes, optou-se por selecionar apenas as 3902 patentes com maior

conectividade da rede. Ademais, a fim de melhorar o layout da rede, aplicaram-se medidas de atração igual a 10 e repulsão igual 1.

5.3.3.2 Procedimentos para análise de acoplamento bibliográfico das patentes

Para realizar a análise de acoplamento bibliográfico, as patentes do campo PN foram organizadas em coluna e as patentes do campo CP (patentes citadas) foram alocadas na linha subjacente às patentes citantes, como ilustrado na Figura 12.

Figura 12 – Organização das patentes citantes e suas referências em colunas

	A	B	C	D	E	F	G		A	B	C	D	E	F	G
1	WO2019015000	EP2494865	WO2003051379	WO2014152494				1	WO2019015000	WO2019015000	WO2018121000	WO2018121000	WO2018121000	WO2018121000	WO2018121000
2	WO2019025384	CN103374574	CN104745531	CN104745532	CN104745534	CN103468820	CN103374575	2	EP2494865	CN103374	WO2011010000	WO2014100000	US2016021000	WO2013100000	CN104829
3	WO2018208728	WO2011058052	WO2014076137	WO2016138525	WO201618346	WO2006028786	WO2006059141	3	WO2003010000	CN104745	WO2014010000	WO2017010000	WO2017010000	WO2015110000	CN105153
4	WO2018195595	WO2014170480	WO2017044776	WO2016152822	US4373071	US4458066	US4500707	4	WO2014100000	CN104745	WO2016100000	WO2014100000	WO2017100000	WO2016100000	WO2015100000
5	WO2018195545	US2016020824	WO2017050105	WO2017184768				5		CN104745	WO2016100000	US4373071		WO2017010000	WO2016010000
6	WO2018132783	WO2013158232	WO2015164740	WO2016142532	WO20170796	WO2016183041	GB2211904	6							
7	WO2018121712	CN104829733	CN105153316	WO2015142675	WO20160476	WO2016073795	WO2016094304	7							
8	WO2018121263	US20160243175	WO2014130540	WO2016149449	US2016007120	US8268798	US8283461	8							
9	WO2018102795	US2006003481	US2016003444	US9422547	WO2015173111	WO2015184228	WO2006070119	9							
10	WO2018073391	WO2014191129	WO2015075195	WO2016134765	US4682199	US5688358	US5682322	10							
11	WO2018064640	US2009008676	US20150151298	US20150267251	US201502980	WO2015103339	WO2014047561	11							
12	CN107723275	CN106103475	CN106544321	CN107243318	CN107723271	WO2017033969	CN106103475	12							
13	WO2018035423	US4647569	US5972701	US20150335705	US2016013800	WO2016036754	WO2016044605	13							
14	WO2018035388	US2003005399	US20030100701	US2005012039E	US201010405E	US2014035979E	US2015003066E	14							
15	WO2018005873	EP264166	EP1519714	EP1664316	EP1766035	EP1781553	US4186183	15							
16	WO2017223330	US2008003186E	US20130245034	WO2015195621	WO20121704	US2008003186E	WO2015195621	16							
17	WO2017191274	WO2009030481	WO2010037539	WO201026641	WO201030332	WO2010281980	WO2012195193	17							
18	WO2017186718	WO2010089290	WO2014191518	WO2014059173	CN103002913	CN104203982	EP1674479	18							
19	WO2017189308							19							
20	WO2017173353	CN1038674	CN101058891	CN103958667	CN104011071	CN104694575	CN106163847	20							

Fonte: elaborado pela autora.

Para o cálculo de acoplamento entre as patentes, bem como para a identificação da intensidade de acoplamento entre as patentes, foi utilizado o código de Castanha (2022), que calcula o acoplamento bibliográfico via software R. O código executa o cálculo pareado entre todos citantes (citante-citante) identificando o tamanho da lista de referências, a quantidade de itens citados em comum (intensidade de acoplamento) e suas respectivas normalizações via Cosseno de Salton e Índice de Jaccard.

Após atender as diretrizes, obtém-se um documento em arquivo txt. com os seguintes elementos: patente 1 a ser acoplada (X1), patente 2 a ser acoplada (X2), número de referências de X1 (refs_x1); número de referências de X2 (refs_x2), quantidade de referências em comum (AB), Saltons_Cossine: Normalização por cosseno de Salton e Jaccard Index: Normalização por índice de Jaccard (Figura 13).

Figura 13 – Recorte dos dados gerados no software R para elaborar a rede de AB

	X1	X2	refs_X1	refs_X2	ABA	Saltons_Cosine	Jaccard_Index
1	CN105154473	CN104920227	8	6	0	0.000000000	0.000000000
2	CN105267240	WO2015191693	4	39	0	0.000000000	0.000000000
3	CN105267240	WO2016004875	4	11	0	0.000000000	0.000000000
4	CN105267240	CN104920227	4	6	0	0.000000000	0.000000000
5	CN105267240	WO2016004043	4	18	0	0.000000000	0.000000000
6	CN105267240	CN105154473	4	8	0	0.000000000	0.000000000
7	CN105331586	WO2016021972	6	22	0	0.000000000	0.000000000
8	CN105331586	US2016024474	6	6	0	0.000000000	0.000000000
9	CN105331586	WO2015191693	6	39	0	0.000000000	0.000000000
10	CN105331586	CN105154473	6	8	0	0.000000000	0.000000000
11	CN105331586	WO2016016894	6	80	0	0.000000000	0.000000000
12	CN105331586	WO2016019144	6	200	0	0.000000000	0.000000000
13	CN105331586	WO2016004875	6	11	0	0.000000000	0.000000000
14	CN105331586	CN104920227	6	6	0	0.000000000	0.000000000
15	CN105331586	CN105267240	6	4	0	0.000000000	0.000000000
16	CN105331586	WO2016004043	6	18	0	0.000000000	0.000000000
17	CN105331586	WO2016014576	6	970	0	0.000000000	0.000000000
18	CN105384825	US2016024474	84	6	0	0.000000000	0.000000000
19	CN105384825	CN105267240	84	4	0	0.000000000	0.000000000
20	CN105384825	EP2985285	84	7	0	0.000000000	0.000000000
21	CN105384825	EP2990416	84	111	0	0.000000000	0.000000000
22	CN105384825	WO2016016894	84	80	0	0.000000000	0.000000000
23	CN105384825	WO2016025454	84	9	1	0.036369648	0.0108695652
24	CN105384825	WO2016028896	84	972	7	0.024497697	0.0066730219
25	CN105384825	WO2016004043	84	18	0	0.000000000	0.000000000
26	CN105384825	CN104920227	84	6	0	0.000000000	0.000000000
27	CN105384825	WO2016014576	84	970	10	0.035032771	0.0095785441
28	CN105384825	WO2016021972	84	22	0	0.000000000	0.000000000
29	CN105384825	WO2016025759	84	106	0	0.000000000	0.000000000
30	CN105384825	WO2016004875	84	11	0	0.000000000	0.000000000

Fonte: dados de pesquisa.

A partir desse resultado, os dados foram transferidos para um arquivo Excel onde foram excluídas a primeira coluna, a coluna refs_x1 e refs_x2 e as colunas Saltons_Cossine e Jaccard_Index, para gerar arquivo .net utilizado, na sequência, no software VOSviewer (CTWS, 2021). Para criar os vértices/nós com as 144 patentes, as seguintes etapas foram realizadas (Figura 14).

- 1) Copiou-se a coluna X1 e removeram-se as patentes duplicadas, obtendo-se as 144 patentes (PN), que foram enumeradas de 1 a 144. Aqui, também se substituiu o cabeçalho X1 por Refs.
- 2) Duplicou-se a coluna Refs., incluindo “aspas” às patentes.
- 3) Criou-se a coluna *Vértices/links 144 reunindo na mesma célula os números e as patentes com aspas. Aqui aplicou-se a fórmula =M2&" "&N2 (M2 e N2 as posições dos elementos na célula).
- 4) Para criar *Edges/nós, primeiro aplicou-se a função =PROCV incluindo os argumentos: valor procurado (X1); matriz da planilha (L:M); número_índices_coluna (2;0) para as duas colunas A e B e depois os elementos foram reunidos utilizando-se a fórmula: =F2&" "&G2&" "&H2

Figura 14 – Recorte da planilha contendo os nós e as ligações (links) para criar o arquivo .net

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	X1	X2	ABA			X1	X2	ABA	*Edges			Refs			*Vertices 144
2	WO2016028896	WO2016014576	462			38	30	462	38 30 462			CN104920227	1	"CN104920227"	1 "CN104920227"
3	WO2016164731	WO2016014576	400			78	30	400	78 30 400			CN105154473	2	"CN105154473"	2 "CN105154473"
4	WO2016164731	WO2016028896	390			78	38	390	78 38 390			CN105267240	3	"CN105267240"	3 "CN105267240"
5	WO2016126608	WO2016014576	310			64	30	310	64 30 310			CN105331586	4	"CN105331586"	4 "CN105331586"
6	WO2016126608	WO2016028896	305			64	38	305	64 38 305			CN105384825	5	"CN105384825"	5 "CN105384825"
7	WO2016164731	WO2016126608	284			78	64	284	78 64 284			CN105477016	6	"CN105477016"	6 "CN105477016"
8	WO2017149515	WO2016028896	251			122	38	251	122 38 251			CN105543313	7	"CN105543313"	7 "CN105543313"
9	WO2017149515	WO2016014576	241			122	30	241	122 30 241			CN105647968	8	"CN105647968"	8 "CN105647968"
10	WO2017149515	WO2016164731	217			122	78	217	122 78 217			CN105748343	9	"CN105748343"	9 "CN105748343"
11	WO2017149515	WO2016126608	164			122	64	164	122 64 164			CN105769740	10	"CN105769740"	10 "CN105769740"
12	WO2017109757	WO2017072590	145			118	109	145	118 109 145			CN106267341	11	"CN106267341"	11 "CN106267341"
13	WO2017072590	WO2016135559	131			109	67	131	109 67 131			CN106282107	12	"CN106282107"	12 "CN106282107"
14	WO2017109757	WO2016135559	131			118	67	131	118 67 131			CN106367386	13	"CN106367386"	13 "CN106367386"
15	WO2016106236	EP3009511	119			60	18	119	60 18 119			CN106474155	14	"CN106474155"	14 "CN106474155"
16	WO2017078807	WO2017079673	118			112	113	118	112 113 118			CN107723275	15	"CN107723275"	15 "CN107723275"
17	WO2016106236	WO2016094880	109			60	58	109	60 58 109			EP2985285	16	"EP2985285"	16 "EP2985285"
18	WO2018005873	EP3009511	102			131	18	102	131 18 102			EP2990416	17	"EP2990416"	17 "EP2990416"

Fonte: elaborado pela autora.

Nessa etapa, os elementos *Vértices 144 e *Edges foram transferidos para arquivo .txt. O arquivo foi executado no software VOSviewer (criar mapa baseado em dados de redes e, posteriormente, escolher a aba pajek, campo arquivo de rede pajek necessário). A partir disso, foi criada a rede de acoplamento bibliográfico de patentes.

5.3.3.3 Procedimentos para análise de cocitação a partir das referências de documentos não patente

Em relação à literatura científica, compreende-se que há pilares bem desenvolvidos a respeito dos métodos de cocitação e acoplamento bibliográfico. Small (1973) apresentou uma forma de analisar as relações entre dois documentos e o nomeou cocitação. Na sequência, White (2001) definiu a imagem de citação como o conjunto de todos os autores com os quais alguém foi cocitado. Portanto, a imagem de citação é construída a partir da literatura publicada pelos pesquisadores da comunidade científica ao construir o referencial teórico-metodológico no desenvolvimento de suas pesquisas. Grácio (2020) reitera as ideias do autor, ao considerar que a proximidade entre os documentos cocitados não é definida pelos próprios autores dos documentos, mas pela comunidade científica que fez uso das ideias e conteúdo dos documentos citados, gerando novos conteúdos.

No processo de desenvolvimento da análise de cocitação da literatura não patente, realizou-se a limpeza da lista de referências, excluindo o que não correspondia à literatura científica ou material bibliográfico, como por exemplo a

“Declaration of Steven R. Trybus in EPO opposition proceedings concerning European Patent No. 3 009 511 dated Sep. 14, 2018, 13 pages”, patentes e lista de escritórios de patentes citados no campo CR (citação a referências). Para cada referência com informação insuficiente foram feitas pesquisas na *National Library of Medicine*¹⁴ e google acadêmico, a fim de recuperar as informações de autoria e demais informações que compõem uma referência, como no exemplo abaixo, em que as informações contidas em preto eram substituídas pelas em vermelho. Do total de 16.237 itens (considerando patentes e itens não identificados) citados, após a limpeza, obtiveram-se 10698 referências citadas.

Figura 15 – Exemplo de erros identificados e que foram substituídos pela referência adequada, seguindo a norma da APA

"Human Microbiome Project Consortium 2012, Structure, function, and diversity of the healthy human microbiome", NATURE, vol. 486, no. 7402, 2012, pages 207 – 14
HUTTENHOWER, C.; GEVERS, D.; KNIGHT, R.; ABUBUCKER, S. Human Microbiome Project Consortium 2012, Structure, function, and diversity of the healthy human microbiome. NATURE, v. 486, n. 7402, 2012, p. 207-214.

Fonte: dados de pesquisa

Os dados foram abertos em csv. para a criação do tesouro de autor, permitindo a padronização dos nomes. Para o instrumento, organizou-se a primeira coluna com os nomes dos autores conforme a lista do VosViewer, com o cabeçalho intitulado *Label* (rótulo – termo errado) e uma coluna denominada *Replace by* (substituir por) em que foram incluídos os nomes com a grafia correta. Esse processo de normalização dos nomes é importante, pois impede que o mesmo autor apareça em nós diferentes na rede.

O software VosViewer permite trabalhar com diversos indicadores de produtividade e relacionais. Van Eck e Waltman (2020) esclarecem que ao trabalhar com as relações de coautoria, citação ou acoplamento bibliográfico, o atributo “Citação” indica o número de citações recebidas por um documento ou o número total de citações recebidas por todos os documentos publicados por uma fonte, que pode ser um autor, uma organização ou um país. Eles também enfatizam que quando se analisa as relações de cocitação, o atributo “Citação” indica o número de citações feitas para uma referência citada, uma fonte citada ou um autor citado. Fundamentado

¹⁴ A National Library of Medicine é uma das mais importantes bibliotecas da área da saúde, sendo operada pelo governo federal dos Estados Unidos. National Library of Medicine (<https://www.ncbi.nlm.nih.gov/>)

no manual elaborado pelos autores, compreende-se que é possível aplicar as técnicas de Cocitação e Acoplamento bibliográfico para analisar as patentes e a literatura científica nelas citadas.

A partir dos dados exportados da base de dados DII, selecionaram-se apenas os campos AU – inventores e o campo CR – referências citadas. Para desenvolver a rede, utilizou-se o software VosViwer. Na primeira etapa do processo, no próprio software, realizou-se a extração em formato txt. das listas de referências citadas, tendo como método a função de análise de cocitação de referências. Após essa etapa, realizou-se um tesouro de referência a partir dos erros encontrados na lista, tais como nomes com registros errados ou referências com duplicidade por conter um sinal que a diferencia da outra referência, obtendo 8233 referências.

A fim de tornar a rede visualmente compreensível, diminuindo o abarrotamento de informação em seu layout, adotou-se como critério que as referências (*nodes* - nós) tivessem no mínimo 10 relações (*links* – laços) para aparecer na rede. Ainda, foi utilizado o critério de atração 10 e repulsão -1 no layout do software. Obtiveram-se 5 clusters e 136 itens (referências mais citadas).

Como Grácio (2020; 2018) menciona a análise de cocitação de autores, em geral, é feita de forma não valorada, ou seja, sem a contagem de frequência de citação, mostrando apenas a presença ou a ausência dos autores que compõem a identidade de citação do citante. Como pretendemos identificar os domínios científicos que compõem a base tecnológica das patentes em células-tronco, esta etapa da pesquisa se detém em apresentar os resultados não valorados, mantendo o foco nas relações temáticas dos autores citados.

5.3.3.4 Procedimentos para análise de acoplamento bibliográfico a partir das referências de documentos não patente

Esta seção apresenta os procedimentos utilizados para obter os resultados que foram apresentados na seção 6.3 do capítulo 6 – Análise e discussão dos resultados.

Diferentemente da pesquisa de Grácio (2020) que utilizou como instrumento para obtenção das informações dos pesquisadores do Currículo Lattes, nesta pesquisa realizou-se a busca pelo ORCID¹⁵ (*Open Researcher and Contributor ID* –

¹⁵ O ORCID é um código de identificação usado para localizar cientistas e outros autores acadêmicos e colaboradores e que o diferencia de todos os outros pesquisadores. ORCID (<https://orcid.org/>)

em português é ID Aberto de Pesquisador e Contribuidor). Embora a ferramenta seja importante para identificar os pesquisadores, tivemos algumas limitações que são mencionadas ao final da seção 5.4. Desta forma, para se obter as informações necessárias, realizou-se uma pesquisa exaustiva, concomitante, pelo Google e pelo ORCID, utilizando-se o nome dos pesquisadores e as palavras-chaves identificadas nas patentes das quais são inventores. Cabe mencionar que todo esse processo é importante para que possamos identificar os temas e compreender as proximidades teórico-metodológicas identificadas nos cluster das redes de acoplamento bibliográfico.

A visualização dos domínios sobre o tema células-tronco, a partir de suas proximidades teóricas, foi realizada por meio da designação da identidade de citação de cada um dos inventores que compõe as 144 patentes. Para isso, os campos inventores (AU) e referências citadas (CR) foram organizados em uma nova planilha Excel, permitindo identificar os referentes teóricos que embasam as invenções e que são citados tanto pelos inventores como pelos examinadores.

Para o desenvolvimento da rede de acoplamento bibliográfico, também se utilizou um tesouro a fim de padronizar os nomes dos pesquisadores. Utilizou-se o software VosViewer para a construção da rede de acoplamento bibliográfico, tendo como critério no mínimo 2 documentos por autor, atração 4 e repulsão -1. Obtiveram-se 15 clusters compostos por um total de 85 pesquisadores.

5.4 LIMITAÇÃO DO ESTUDO

Nesta pesquisa, as limitações dizem respeito aos dados registrados na base de dados, uma vez que a DII apresenta registros com informações incompletas, possui problema na padronização dos nomes dos inventores e dos depositantes (titulares das patentes). Ademais, a compreensão dos dados é complexa, visto que uma única patente apresenta uma variedade de informações. Essas limitações também foram relatadas no estudo de Ávila-Robinson e Sengoku (2017) que apontam a necessidade de uma limpeza e triagem exaustiva para um bom resultado da pesquisa.

Observa-se que o maior problema com os dados de patente para o desenvolvimento de uma pesquisa voltada à análise das citações, é a falta de padronização das informações. A quantidade de informações presentes em uma patente faz com que não exista um cuidado com os elementos básicos que compõem

uma referência, por exemplo. Nas listas de referências de documentos não patente, existem muitas informações fundamentais incompletas, como o nome do autor da obra, o ano da publicação ou a fonte onde aquela obra foi consultada. Nesse campo, encontramos problemas como o ilustrado na Figura 15. Além disso, observou-se que, embora existam campos para incluir citações às patentes (CP) e campo exclusivo para outras referências (CR), não há um cuidado para incluir outras referências que não patentes. Citam-se ainda, erros ocasionados por documentos cujo idioma está em japonês, chinês, entre outros.

Figura 16 – Ilustração dos problemas encontrados nos documentos de patentes pela falta de padronização das informações

"Advance STEM Murine Embryonic Stem Cell Culture Products,” ThermoScientific, page 27, 2010/2011.
"ANIMAL CELL CULTURE", 1987
ACS Med. Chem. Lett., 2010, 1, 39-43
Adames <i>et al.</i> , 1985, Nature 318:533-538
Accession No. NM_001772.3. accessed Oct. 13, 2015 from http://www.ncbi.nlm.nih.gov/nucore/130979980
See also references of EP 3189828A4

Fonte: dados da pesquisa.

Na Figura 16, observamos que há referência a um artigo da revista Nature, em que incluem apenas o sobrenome do primeiro autor, mas não seu prenome ou suas iniciais. Além disso, não há uma padronização dos autores, o que dificulta os estudos que usam, na metodologia, indicadores de citação uni-variadas ou relacionais.

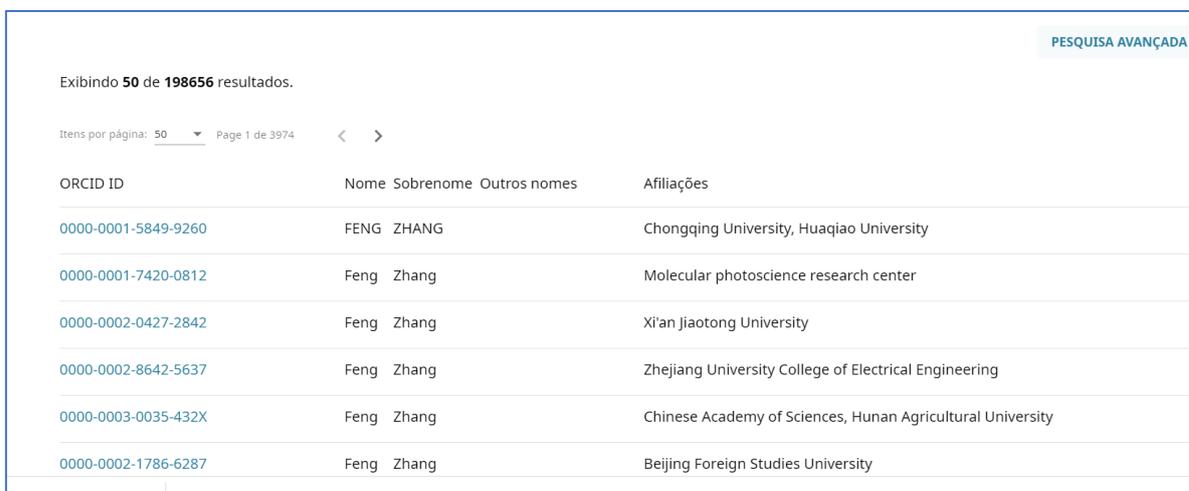
Também foi possível observar que, embora a base de dados informe separadamente as referências incluídas pelo examinador e aquelas realizadas pelo inventor, no momento da exportação dos dados, essas informações aparecem juntas, sem distinção entre quem (inventor ou examinador) inseriu qual referência. Esse fato impede a realização de estudos que analisem o comportamento informacional desses dois tipos de atores no ambiente tecnológico.

Embora existam outras bases de dados de patentes, tais como Orbit Questel e Patentscope, que possibilitam o uso de técnicas patentométricas, entende-se que para este estudo a escolha da Base de Dados DII é a melhor opção, por apresentar uma boa cobertura, os campos de citação tanto às patentes como aos documentos

não-patentes e por ser uma base de dados disponibilizada pela Comunidade Acadêmica Federada (CAPES) às Universidade públicas.

Durante o processo de busca os pesquisadores/inventores no ORCID, um fator limitador foi a falta de informações suficientes para a pesquisa como, por exemplo, as palavras-chave. Alguns pesquisadores não têm como hábito incluir termos (palavras-chave) que os relacionem às suas pesquisas de ponta, o que dificulta o processo de busca no ORCID. Como exemplo, em relação ao assunto CRISPR, citamos o pesquisador Feng Zhang, que buscando, no ORCID, pelo nome, obtivemos 198656 resultados (Figura 17). Quando buscamos esse pesquisador em pesquisa avançada, utilizando o nome e o termo CRISPR, nenhum pesquisador é revocado. Ao incluir o nome e o termo Stem-cell, recuperamos 2 pesquisadores, mas nenhum deles é o pesquisador deste estudo. Ao incluir seu nome e uma das afiliações conhecidas do pesquisador, *Stanford University*, identificamos precisamente o autor. Isso demonstra a importância do preenchimento adequado dos campos de um instrumento relevante para a ciência.

Figura 17 – Processo de busca realizado no ORCID



The screenshot shows a search results page for 'Feng Zhang' on ORCID. It displays 50 results out of 198,656. The results are listed in a table with columns for ORCID ID, Name (First, Last, Other), and Affiliations. The first few results are as follows:

ORCID ID	Nome Sobrenome Outros nomes	Afiliações
0000-0001-5849-9260	FENG ZHANG	Chongqing University, Huaqiao University
0000-0001-7420-0812	Feng Zhang	Molecular photoscience research center
0000-0002-0427-2842	Feng Zhang	Xi'an Jiaotong University
0000-0002-8642-5637	Feng Zhang	Zhejiang University College of Electrical Engineering
0000-0003-0035-432X	Feng Zhang	Chinese Academy of Sciences, Hunan Agricultural University
0000-0002-1786-6287	Feng Zhang	Beijing Foreign Studies University

Fonte: Print da tela do computador.

5.5 PLANO DE GESTÃO DE DADOS

No cenário atual das pesquisas científicas, onde boas pesquisas são produzidas com grande quantidade de dados, há uma preocupação em otimizar o tempo do pesquisador, evitando que esses façam coletas anteriormente realizadas.

Para isso, pesquisadores como Sayão e Sales (2015) vêm discutindo e planejando métodos de preservação desses dados. Borgman (2010) reforça que, para que os dados estejam acessíveis, deve haver por parte dos pesquisadores o interesse em tornar esses dados disponíveis e compartilháveis. Bertin, Visoli e Drucker (2017) dizem que a alta capacidade das instituições (universidades, centros de pesquisa e organizações) de gerar dados de pesquisa, fez com que a preocupação com o seu gerenciamento aumentasse. Para eles, a Gestão de Dados de Pesquisa se tornou um elemento propulsor para o avanço da ciência e a tecnologia no que tange a *e-science*¹⁶. Sendo assim, o plano de gestão de dados deste estudo segue o checklist proposto pelo *Inter-University Consortium for Political and Social Research* ([2012]) e por Sayão e Sales (2015).

Os dados brutos utilizados no estudo serão disponibilizados, transcorrido um ano da defesa e publicação da tese, em uma página virtual criada para consulta de quaisquer interessados, respeitando os princípios éticos da pesquisa.

¹⁶ Refere-se a um movimento que busca a concepção de uma ciência atuante no uso intensivo de dados e na colaboração por meio do uso de plataformas que contribuem com essa pesquisa mais colaborativa.

6 ANÁLISE E DISCUSSÃO DOS RESULTADOS

Neste capítulo, encontram-se os resultados relativos à análise da contribuição dos métodos de cocitação e de acoplamento bibliográfico para a caracterização do domínio tecnológico no âmbito das células-tronco.

O Manual de Estatística de Patentes da OCDE (2009) alerta para o fato de que trabalhar com documentos de patentes é um processo complexo que exige conhecimento das informações ali contidas. Ao estabelecer os seis passos para analisar a proximidade entre autores, McCain (1990) também aponta para a importância de se conhecer o campo científico analisado e sinaliza para a necessidade de se buscar a contribuição dos especialistas da área para uma interpretação mais fidedigna dos resultados.

Partindo desse entendimento, em muitos momentos, ao longo das análises realizadas na presente pesquisa, fez-se uso do DeCS/MeSH¹⁷ - Descritores em Ciências da Saúde - criado pela BIREME, que contribui para se obter uma linguagem única (vocabulário controlado) na atividade de indexação de artigos científicos e demais documentos na área da saúde.

A patente que apresentou o maior número de citações, compondo o corpus da pesquisa, foi EP3009511, com 297 citações, também apresentou em sua lista de referências 681 patentes citadas e 1104 referências de outros tipos de documentos (Apêndice B). Ela trata sobre novas enzimas do sistema CRISPR - sequências de ácidos nucleicos repetitivas, que funcionam como um sistema de defesa antiviral. A invenção fornece sistemas, métodos e composições para direcionar ácidos nucleicos. Além disso, essa é uma das mais modernas técnicas de edição genética aplicada em diversos organismos, entre eles plantas, contribuindo para o melhoramento genético na agricultura, caso da patente mencionada. Tem como primeiro inventor o bioengenheiro sino-americano Feng Zhang¹⁸, especialista em optogenética¹⁹ e CRISPR, integrante da Massachusetts Institute of Technology e do Broad Institute. Entre as 10 patentes do corpus da pesquisa que receberam maior número de citações,

¹⁷ Esse instrumento foi elaborado a partir do MeSH – Medical Subject Headings da U.S. National Library of Medicine (NLM) com o objetivo de permitir o uso de terminologia comum para pesquisa em múltiplos idiomas, proporcionando um meio consistente e único para a recuperação da informação. DeCS/MeSH (<https://decs.bvsalud.org/>).

¹⁸ Informações fornecidas no site da Beam Therapeutics (<https://beamtx.com/team/feng-zhang/>)

¹⁹ Combinação de métodos genéticos e ópticos no controle de eventos específicos com precisão temporal em células alvo fe um sistema biológico intacto e funcional.

outras duas (WO2016205613, WO2016106236) têm como inventor o pesquisador Feng Zhang e tratam sobre a mesma temática. Além dessas, outra que trata sobre CRISPR é a patente (US2016145646) – 59 citações – cuja invenção foi desenvolvida pelo doutor em bioquímica David Frendewey, que atua como diretor Executivo de Descoberta e Triagem da Regeneron Pharmaceuticals (Apêndice B).

Caetano (2018, p.96) afirma que essa técnica foi mencionada pela primeira vez na década de 80 pela pesquisadora “[...] Emmanuelle Charpentier, da Hannover Medical School, na Alemanha, e por [...] Jennifer Doudna, da Universidade de Berkeley na Califórnia [...]”. Contudo, foi Feng Zhang, em 2013, que conseguiu aplicar a técnica em células humanas, sendo concedida, a partir disso, a primeira patente norte-americana sobre a técnica. Além disso, o CRISPR tornou financeiramente mais acessível o processo de reprogramação genética.

Olhar o corpus da pesquisa nos permite compreender as patentes e demais documentos citados. Spinak (2003) afirma que se uma patente é altamente citada há indícios de que seja uma invenção com avanço tecnológico muito relevante. Segundo ainda o autor, a frequência de citação de uma determinada patente indica seu valor comercial.

6.1 CARACTERIZAÇÃO DO REFERENCIAL TECNOLÓGICO PRESENTE NO UNIVERSO DAS PATENTES

Conforme mencionado na introdução desta pesquisa, os estudos patentométricos contribuem para medir a produtividade, o impacto das tecnologias nas comunidades (social e científica), além de identificar a elite tecnológica e contribuir para o processo de gestão da propriedade intelectual das empresas e instituições, que têm como atividade o desenvolvimento tecnológico e a prática do patenteamento, além de outros indicadores como o valor econômico da patente.

O Manual para Estatísticas de Patentes da OCDE (2009) afirma que as citações às patentes permitem rastrear os efeitos da propagação do conhecimento de uma tecnologia, bem como aferir a curva de obsolescência das tecnologias. Compreendendo a importância da pesquisa e dos avanços tecnológicos em células-tronco para o desenvolvimento da medicina regenerativa, engenharia genética, biomedicina e demais campos, o Apêndice B apresenta o corpus desse estudo

(ordenado por citações recebidas), referente às patentes mais citadas no período de 2016 a 2020.

Nesse sentido, observa-se (Tabela 2) que entre as 4915 patentes citadas no corpus da pesquisa, a patente mais citada (US8697359, com 35 citações) fornece sistemas, métodos e composições para alterar a expressão de sequências de genes alvo e produtos de genes relacionados e tem como único inventor o pesquisador Feng Zhang.

O número total de patentes citadas nas 144 patentes analisadas para análise dessa pesquisa foi de 4915 patentes. A Tabela 2 apresenta as 88 patentes mais citadas pelas 144 patentes que compõem o universo da pesquisa, resultante da aplicação da Lei do elitismo de Price no conjunto das 4915 patentes citadas.

Tabela 2 – Conjunto das 88 patentes mais citadas no corpus com suas respectivas contagens de citações

Patentes Citadas				Patentes Citadas			
#		Quant.	Freq.	#		Quant.	Freq.
1	US8697359	35	0,27	45	WO2013074916	13	0,10
2	WO2013176772	32	0,24	46	WO2016025385	13	0,10
3	US20140068797	27	0,21	47	WO2014204725	13	0,10
4	WO2014093622	24	0,18	48	WO2012099973	13	0,10
5	WO2014134165	21	0,16	49	WO2014099750	13	0,10
6	WO2013019615	21	0,16	50	US8911993	13	0,10
7	WO2013123061	20	0,15	51	US8865406	13	0,10
8	US8399645	20	0,15	52	US20140050708	13	0,10
9	US20140099309	20	0,15	53	US20140310828	13	0,10
10	WO2013126712	19	0,15	54	US20140179006	13	0,10
11	EP3009511	19	0,15	55	US20140273234	13	0,10
12	WO2014065596	18	0,14	56	US20140227787	13	0,10
13	WO2012079000	18	0,14	57	US20100076057	13	0,10
14	WO2015142675	18	0,14	58	WO2012082841	12	0,09
15	WO2014093661	18	0,14	59	WO2014153270	12	0,09
16	WO2015090229	17	0,13	60	US6352694	12	0,09
17	WO2015089419	17	0,13	61	US7175843	12	0,09
18	WO2013126794	17	0,13	62	US7067318	12	0,09
19	WO2014093712	17	0,13	63	US6534055	12	0,09
20	US7446190	17	0,13	64	WO2005047334	12	0,09
21	US20140179770	17	0,13	65	US6692964	12	0,09
22	WO2016028896	16	0,12	66	US6905874	12	0,09
23	WO2014165825	16	0,12	67	US6797514	12	0,09
24	WO2010003118	16	0,12	68	US7144575	12	0,09
25	WO2016014565	15	0,11	69	US6867041	12	0,09
26	WO2015048577	15	0,11	70	US7232566	12	0,09
27	WO2014093595	15	0,11	71	US6887466	12	0,09
28	US20150183881	15	0,11	72	US5858358	12	0,09
29	WO2014055442	14	0,11	73	US8389282	12	0,09
30	WO2014018423	14	0,11	74	US5883223	12	0,09
31	WO2013173820	14	0,11	75	US6905680	12	0,09
32	WO2014055657	14	0,11	76	US6905681	12	0,09
33	WO2014145252	14	0,11	77	US7172869	12	0,09
34	WO2014204724	14	0,11	78	US20140248702	12	0,09
35	WO2014130635	14	0,11	79	US20140199767	12	0,09

36	US5906936	14	0,11	80	US20140113370	12	0,09
37	US7741465	14	0,11	81	US20140273231	12	0,09
38	US20140242699	14	0,11	82	US20140242700	12	0,09
39	US20140315985	14	0,11	83	US20140186843	12	0,09
40	US20140273232	14	0,11	84	US20150353642	12	0,09
41	US20150190506	14	0,11	85	US20140234972	12	0,09
42	US20140242664	14	0,11	86	US20140357530	12	0,09
43	US20140186919	14	0,11	87	US20120321667	12	0,09
44	WO2016014576	13	0,10	88	US20060121005	12	0,09
				#	#	763	5,84
				#	#	12300	94,16
Total						13063	100

Fonte: elaborado pela autora.

O que se observa é que as patentes mais citadas, tratam sobre métodos e composições para modificação genética. A segunda patente mais citada WO2013176772 (32 citações) fornece métodos de modulação da transcrição de um ácido nucleico alvo em uma célula alvo, geralmente envolvendo o contato do ácido nucleico alvo com um polipeptídeo²⁰ Cas9 (genes associados ao CRISPR). Por sua vez, a patente US20140068797 (27 citações) e tem como primeira inventora uma das pesquisadoras pioneiras sobre CRISPR, a doutora Jennifer Doudna²¹. Ela, juntamente com a pesquisadora Emmanuelle Charpentier foram laureadas pelo desenvolvimento da técnica de edição de genoma, recebendo o Nobel de Química em 2020. A quarta patente com mais citações (WO2014093622) também pertence ao pesquisador, Feng Zhang, que possui a patente mais citada no corpus da pesquisa.

Com 21 citações, a quinta patente mais citada (WO2014134165) trata sobre uma invenção que fornece métodos e composições para aumentar a resposta imune contra cânceres e patógenos e tem como inventores Renier Brentjens e Hollie Pegram, ambos especialistas em leucemia de célula T (doença maligna dos linfócitos T na medula óssea, timo ou sangue). A patente WO2013019615 (21 citações) trata de uma invenção sobre um receptor para troca de proteína quimérica que promove um aumento da resposta imune. Essa invenção tem como pesquisador responsável os imunologistas Carl H June e Yangbing Zhao, ambos professores da Universidade da Pensilvânia. A patente WO2013123061 (20 citações) tem como inventor o pesquisador e professor em hematologia-oncológica, Dr. Michael Claus V Jensen que criou um receptor de antígeno quimérico e uso terapêutico para atingir antígenos únicos em células cancerígenas e a patente US8399645 (20 citações) que se refere à

²⁰ Os polipeptídeos são compostos aproximadamente de 13 ou mais aminoácidos. As PROTEÍNAS são polipeptídeos lineares geralmente sintetizados nos RIBOSSOMOS.

²¹ Britannica.com (<https://www.britannica.com/biography/Jennifer-Doudna>)

invenção de um receptor quimérico que tem como objetivo a ativação e expansão de células para usos terapêuticos. Os inventores são Dario Campana pesquisador do departamento de Oncologia e Patologia do Hospital de Pesquisa Infantil St Jude e Chihaya Imai, professor pediatra e diretor adjunto na Universidade de Niigata.

O que se observa nas 10 patentes mais citadas entre as 4915 citadas é que as 5 primeiras estão mais voltadas à temática CRISPR, uma técnica mais simples para originar células-tronco e fornecem informações importantes sobre o processo de reprogramação celular (NAMBIAR *et al.*, 2022), indicando um forte interesse no ambiente científico e tecnológico. Por sua vez, as outras 5 patentes mencionadas acima tratam de receptores que buscam aumentar a resposta imunológica e contribuir no tratamento de doenças neoplásicas.

A fim de aferir se as temáticas apresentadas nas patentes mais citadas também estão presentes entre as menos citadas, analisou-se alguns documentos, buscando de forma aleatória as patentes que receberam uma citação. Entre as 4915 patentes citadas, 2693 (54,79%) foram citadas apenas uma vez.

A patente WO2007101347, com uma citação (0,01%) é uma invenção que tem como objetivo regular os distúrbios proliferativos ou apoptose²² desregulada das células, que causam doenças autoimunes, distúrbios inflamatórios e até câncer. A patente CN103002913A, também com uma citação (0,01%), trata de uma composição ou método para tratamento do Lúpus Eritematoso Sistêmico²³ e da síndrome de Sjögren²⁴. A patente EP1669441A1 (0,01%) refere-se a tecnologia para diferenciar células mesenquimais em hepatócitos, relativa à medicina regenerativa do fígado para animais e humanos com lesões no fígado. Por sua vez, a invenção representada pelo número de patente WO2012090104, com uma citação, voltada para a indústria farmacêutica, refere a métodos e composições para projetar novas terapias relacionando drogas. A patente US6987835 (0,01%), refere-se a um tubo de raios X em miniatura em uma sonda, para uso em lúmens estreitos, como vasos sanguíneos humanos, ou com um aplicador em cavidades naturais ou cirúrgicas. Observa-se que

²² Mecanismo regulado de morte celular caracterizado por alterações morfológicas distintas no núcleo e no citoplasma, incluindo a clivagem endonucleolítica do DNA genômico. É, quando regulado, um processo natural de morte celular.

²³ doença autoimune, ou seja, que ocorre quando o sistema imunológico do seu corpo ataca seus próprios tecidos e órgãos.

²⁴ Doença autoimune, em que o sistema imunológico do corpo ataca suas próprias células saudáveis que produzem saliva e lágrimas. Geralmente ocorre junto com outras doenças, como a artrite reumatoide e o lúpus.

as patentes que receberam uma citação apresentam uma abordagem mais específicas, sobre temáticas bem delimitadas.

Elucidar essas informações, ainda que de forma estimativa, dada a inviabilidade de se descrever cada patente e seus inventores e cessionários, contribui para o entendimento sobre como o cenário da biotecnologia vem se desenvolvendo no que diz respeito à inovação. É possível observar que as patentes mais citadas no corpus da pesquisa têm, na maioria, pesquisadores americanos ou com algum laço acadêmico-científico com universidades e empresas dos Estados Unidos. Cabe recordar que isso só é possível analisar olhando as afiliações dos inventores e do titular, pois o número da patente apresenta o código onde ela foi protegida e não a localização geográfica dos inventores. Esse resultado está muito relacionado ao investimento que o país faz na pesquisa básica e na formação de quadros de alto nível em pesquisa genética. O Programa de Genoma Humano (PGH), por exemplo, é um consórcio global de pesquisas liderado pelos Estados Unidos desde 1990. Embora tenha a participação de países como França, Alemanha, Japão, Canadá e Brasil, entre outros, apenas ele é responsável por dois terços das pesquisas envolvendo sequenciamento genético, que contou com a contribuição do governo americano através do *National Institutes of Health (NIH)* (CORRÊA, 2002).

Embora a presente pesquisa não tenha como foco uma análise terminológica, considera-se significativo apresentar os inventores e os assuntos tratados em algumas patentes, especialmente as mais citadas, correspondente às de maior impacto no avanço do tema analisado, para entender como se configura o cenário tecnológico envolvendo células-tronco, visto que na análise de domínio Hjørland (2002), ao elencar as 11 abordagens, considera importante relacionar aos estudos métricos, abordagens que tratem de aspectos históricos e epistemológicos.

Ainda que pareça provável que os assuntos das patentes estejam voltados para o tratamento de doenças patogênicas e cancerígenas, vale aqui recordar que a busca pelas patentes desse estudo foi apenas com o termo “Stem cell” e este pode estar relacionado a células-tronco humanas ou vegetais, bem como abranger tratamentos de doenças, como técnicas para combater o envelhecimento, voltado para o cenário da indústria de cosméticos. Como exemplo, identificamos a patente CN105748343 (Apêndice B), que trata de um tipo de composição de extrato de plantas e suas aplicações em cosméticos. A patente tem como inventor o chinês Chen Songbin, presidente da Guangdong Kewei Biotechnology Co. Ltd. O nicho de negócio dessa

empresa inclui pesquisa e desenvolvimento de tecnologia de produtos bioquímicos; serviços técnicos, desenvolvimento de tecnologia.

Essa análise preliminar das patentes que compõem o corpus da pesquisa e suas abordagens contribui para o embasamento das análises seguintes em que as relações de similaridade entre as patentes são tratadas a partir da análise de acoplamento (relação citante-citante) e análise de cocitação (relação citado-citado) de patente.

6.2 CARACTERIZAÇÃO DO REFERENCIAL CIENTÍFICO PRESENTE NO CONJUNTO DAS PATENTES EM CÉLULAS-TRONCO

As análises baseadas em referências estão categorizadas nos indicadores de produção. Sobre esta perspectiva, entende-se que as referências mencionadas por um autor ao longo de sua trajetória acadêmica e de pesquisa discorre muito sobre seus interesses teóricos metodológicos e sua identidade científica. Essa identidade de citação, mencionada por White (2001), torna-se mais evidente quando há reincidência dos autores citados ao longo das publicações. O termo reincidência da referência, mencionado por Grácio (2020), é denominado recitação por White (2001) e pode ocorrer de forma sincrônica - em que o pesquisador citante referência diversas vezes um mesmo autor em uma lista de referência - ou diacrônica, em que o pesquisador cita o mesmo autor em diversas obras publicadas ao longo da sua carreira.

Grácio (2020) afirma que alguns autores são citados uma única vez, em função dos objetivos específicos estabelecidos em uma única pesquisa. Dessa forma, a autora considera que a reincidência de citação de um autor nas produções subsequentes demonstra a relevância que ele tem para a composição das fundamentações científicas do pesquisador.

Entretanto, ainda sobre a perspectiva da literatura científica, algumas questões relacionadas às razões pelas quais um pesquisador cita o trabalho de outro permeiam a ciência, visto que se sabe que tal prática pode ocorrer por apreciação real aos pesquisadores, pelo arcabouço científico ou por critérios volúveis (GARFIELD, 1979). Gilbert (1977) afirmou que os autores optam por citar artigos que reconhecem como oficiais para justificar a validade, novidade e importância de seu próprio trabalho. No entanto, como mencionam Vanz e Caregnato (2003), as análises de citação devem

ser interpretadas com cautela, uma vez que a prática da citação pode ser influenciada por fatores culturais, sociais, políticos e econômicos.

Cabe observar que, para que possamos identificar a identidade de citação dos pesquisadores, cuja análise é desenvolvida na seção 6.3, é preciso que haja recitação diacrônica, ou seja, os citados devem aparecer em pelo menos duas patentes distintas. Nesta etapa da pesquisa, apresentamos um panorama dos pesquisadores mais citados nas 144 patentes. Os dados extraídos da base não permitem identificar se a citação foi feita pelo examinador ou pelo inventor, visto que elas são aglutinadas no campo CR sem nenhuma descrição do citante. Este fato nos leva a refletir sobre a importância desta informação, visto que o examinador tem um papel crucial no processo que torna a invenção uma patente concedida (WORLD INTELLECTUAL PROPERTY ORGANIZATION, 2004).

Como é possível observar na Tabela 3, a referência que recebeu mais citação (74 citações) tem como primeiro autor o pesquisador Martin Jinek, vinculado ao Departamento de Bioquímica da Universidade de Zurique, com pós-doutorado com a pesquisadora Jennifer A. Doudna, premiada com Nobel em Química, uma das precursoras sobre pesquisa em CRISPR, inventora da patente US20140068797 (seção 6.1) e coautora do artigo. Martin Jinek também aparece como coautor (o último autor) no 11º artigo (22 citações), cujo tema aborda a proteína Cas9 que associada ao CRISPR demonstra ser uma técnica de baixo custo muito promissora para tratamento de doenças (CAETANO *et al.*, 2019).

O segundo artigo mais citado (57 citações) tem como primeiro autor Prashant Mali, professor da Universidade da Califórnia, com formação em Engenharia Elétrica, doutor em Engenharia Biomédica. Lidera um laboratório de pesquisa que tem como foco desenvolver ferramentas moleculares para engenharia de genoma, transcriptoma²⁵ e proteoma²⁶ e sua aplicação à interpretação sistemática do genoma, além de aplicações de terapia genética. Ainda, sob autoria de Prashant Mali e coautores, tem-se o quarto artigo mais citado (30 citações) e que tem como coautor o pesquisador George M Church do Departamento de Genética da Escola de Medicina de Harvard e líder da pesquisa sobre Biologia Sintética no Wyss Institute. Em 1984,

²⁵ É o conjunto completo de transcritos (RNAs mensageiros, RNAs ribossômicos, RNAs transportadores e os microRNAs) de um dado organismo.

²⁶ É o conjunto de proteínas e variantes de proteínas que podem ser encontrados numa célula específica quando esta está sujeita a um certo estímulo.

Church desenvolveu o primeiro método de sequenciamento genômico direto, que resultou na primeira sequência do genoma. Ajudou, ainda, a iniciar o Projeto Genoma Humano em 1984 e o Projeto Genoma Pessoal em 2005. Ele aparece como coautor em outras 16 publicações citadas no corpus do estudo, o que indica um grau de relevância significativo deste pesquisador. Ademais, George M. Church é inventor das patentes WO2016089866 e WO2016141137 que tratam, respectivamente, de sistemas guiados por RNA para edição de genes *in vivo* e de métodos geração de tecido humano funcional.

A terceira referência mais citada é um clássico da pesquisa sobre genética e aborda estudos e métodos de clonagem molecular, tem como autor Sambrook e Church. Sambrook foi um biólogo molecular conhecido por seus estudos sobre vírus tumorais de DNA e células normais e neoplásticas.

A quarta referência mais citada (43 citações) é de autoria do Stephen Frank Altschul, matemático americano responsável por desenvolver algoritmos como BLAST²⁷, usados no campo da Bioinformática. Além disso, esse pesquisador tem outras duas publicações citadas no corpus desta pesquisa, sendo o segundo artigo mais citado (24 citações), localizado na 7ª posição da tabela, um artigo abordando sobre as ferramentas amplamente utilizadas para pesquisar bancos de dados de proteínas e DNA em busca de similaridades de sequências.

Tabela 3 – Parcial de referentes teóricos que sustenta o acoplamento bibliográfico dos pesquisadores

#	REFERÊNCIAS CITADAS	CITAÇÕES
1	Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., & Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. <i>Science (New York, N.Y.)</i> , 337(6096), 816–821.	74
2	Mali, P., Yang, L., Esvelt, K. M., Aach, J., Guell, M., DiCarlo, J. E., Norville, J. E., & Church, G. M. (2013). RNA-guided human genome engineering via Cas9. <i>Science (New York, N.Y.)</i> , 339(6121), 823–826.	57
3	Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989). <i>Molecular cloning: a laboratory manual</i> (No. Ed. 2). Cold spring harbor laboratory press.	45
4	Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. <i>Journal of molecular biology</i> , 215(3), 403–410.	43
5	Mali, P., Aach, J., Stranges, P. B., Esvelt, K. M., Moosburner, M., Kosuri, S., Yang, L., & Church, G. M. (2013). CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. <i>Nature biotechnology</i> , 31(9), 833–838.	30

²⁷ BLAST é um algoritmo e programa para comparar informações de sequências biológicas primárias, como sequências de aminoácidos de proteínas ou nucleotídeos de sequências de DNA e/ou RNA .

6	Ran, F. A., Hsu, P. D., Lin, C. Y., Gootenberg, J. S., Konermann, S., Trevino, A. E., Scott, D. A., Inoue, A., Matoba, S., Zhang, Y., & Zhang, F. (2013). Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. <i>Cell</i> , 154(6), 1380–1389.	27
7	Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. <i>Nucleic acids research</i> , 25(17), 3389–3402.	24
8	Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P. D., Wu, X., Jiang, W., Marraffini, L. A., & Zhang, F. (2013). Multiplex genome engineering using CRISPR/Cas systems. <i>Science (New York, N.Y.)</i> , 339(6121), 819–823.	24
9	Brentjens, R. J., Davila, M. L., Riviere, I., Park, J., Wang, X., Cowell, L. G., Bartido, S., Stefanski, J., Taylor, C., Olszewska, M., Borquez-Ojeda, O., Qu, J., Wasielawska, T., He, Q., Bernal, Y., Rijo, I. V., Hedvat, C., Kobos, R., Curran, K., Steinherz, P., ... Sadelain, M. (2013). CD19-targeted T cells rapidly induce molecular remissions in adults with chemotherapy-refractory acute lymphoblastic leukemia. <i>Science translational medicine</i> , 5(177), 177ra38.	23
10	Jiang, W., Bikard, D., Cox, D., Zhang, F., & Marraffini, L. A. (2013). RNA-guided editing of bacterial genomes using CRISPR-Cas systems. <i>Nature biotechnology</i> , 31(3), 233–239.	23
11	Nishimasu, H., Cong, L., Yan, W. X., Ran, F. A., Zetsche, B., Li, Y., Kurabayashi, A., Ishitani, R., Zhang, F., & Nureki, O. (2015). Crystal Structure of <i>Staphylococcus aureus</i> Cas9. <i>Cell</i> , 162(5), 1113–1126.	23
12	Anders, C., Niewoehner, O., Duerst, A., & Jinek, M. (2014). Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. <i>Nature</i> , 513(7519), 569–573.	22
13	Kalos, M., Levine, B. L., Porter, D. L., Katz, S., Grupp, S. A., Bagg, A., & June, C. H. (2011). T cells with chimeric antigen receptors have potent antitumor effects and can establish memory in patients with advanced leukemia. <i>Science translational medicine</i> , 3(95), 95ra73.	21
14	Riechmann, L., Clark, M., Waldmann, H., & Winter, G. (1988). Reshaping human antibodies for therapy. <i>Nature</i> , 332(6162), 323–327.	20
#		456
#	Outras	13.921
	Total de citações	14.377

Fonte: elaborado pela autora.

Com 27 citações, identificamos o artigo da pesquisadora F Ann Ran, colaboradora na Universidade de Harvard e do Instituto McGovern para Pesquisa do Cérebro, Instituto Broad, Instituto de Tecnologia de Massachusetts. Ela contribuiu para a pesquisa sobre temas como: edição de Cas9, Guia de RNA, genoma e CRISPR, o que pode ser observado no sétimo artigo mais citado, em que a pesquisadora colabora com o pesquisador Le Cong. Os autores versam sobre a importância de tecnologias precisas sobre edição genômica e que o sistema CRISPR apresentam tal tecnologia.

A verificação parcial das referências que compõe o objeto deste estudo versão sobre o domínio da edição genômica e sobre a técnica CRISPR, que é um dos termos que frequentemente aparece na análise deste estudo. Contudo, o CRISPR é frequentemente usado como um sinônimo de edição genômica, mas é o acoplamento

do CRISPR e do sistema Cas9 que permite a deleção seletiva do DNA (CAETANO *et al.*; 2019).

As referências aqui apresentadas mostram que estudos que abordam a temática células-tronco (*Stem-cells*) apresentam uma intensa colaboração científica com interlocução com diversas áreas do conhecimento, como as engenharias, ciências biológicas, ciências da saúde e ciências exatas. A colaboração é muito evidente desde 1990, com o início do Projeto Genoma Humano, que teve a contribuição de diversos países. Solnick, Matthews e Ecklund (2011) afirmam que a pesquisa com células-tronco é uma área ideal para o estudo da globalização científica, uma vez que é influenciada por decisões políticas, que se diferem entre fronteiras, e é objeto de muitos diálogos e redes internacionais. No estudo, as autoras analisaram artigos dos EUA e do Reino Unido porque, segundo elas, esses dois países têm ambientes políticos diferentes, mas são líderes em pesquisa com células-tronco. Na pesquisa, elas observaram que, entre as principais nações que colaboraram com os EUA e o Reino Unido, os únicos países que não se sobrepuseram foram a China e a Coreia do Sul (ambas colaborações dos EUA) e a Suíça e a Suécia (ambas colaboradoras do Reino Unido), sendo que Alemanha, Itália, Holanda, França, Espanha, Canadá e Japão colaboraram com ambos os países.

Considerando o número de referências que foi possível analisar neste estudo (8233) e o número de citações distribuídas entre elas, é possível afirmar que há um número muito expressivo de referências que receberam apenas uma citação (6136) e que 2097 referências receberam duas ou mais citações, o que nos leva a considerar que embora as citações sejam importantes, elas são dispersas.

6.3 ACOPLAMENTO DOS INVENTORES EM RELAÇÃO AO REFERENCIAL CIENTÍFICO

Descrito por Kesller (1963), o acoplamento bibliográfico configura a sobreposição das listas referências de dois artigos, com as referências em comum representando a intensidade do acoplamento bibliográfico entre esses dois artigos. Desta forma, é possível afirmar que quanto mais referências dois artigos têm em comum, maior é a intensidade de acoplamento entre eles. Kesller (1965) considera que a coleção de referências mencionadas pelos autores em suas publicações descreve o ambiente intelectual no qual estão inseridos.

A Figura 18 apresenta a rede de acoplamento bibliográfico entre os inventores analisados, conforme a proximidade teórica observada a partir das referências bibliográficas citadas, denominada por White (2001) de identidade de citação.

O primeiro domínio identificado, denominado cluster C1, é formado por pesquisadores e grupo diretivo da empresa Fate Therapeutics, que atua no segmento da biofarmacêutica em estágio clínico. Compreende inventores de ponta sobre imunoterapias celulares derivadas de células-tronco pluripotentes induzidas (iPSC) para pacientes com câncer e distúrbios autoimunes. Valamehr é o Diretor de Pesquisa e Desenvolvimento e vice-presidente de Imunoterapia do Câncer e Raedun Clarke é diretora sênior. Ryan Bjordahl é diretor sênior no setor de Imunoterapia da empresa. Valamehr e Bjordahl apresentam coautoria recorrente em alguns artigos analisados.

O cluster C2 é totalmente voltado para materiais e métodos de edição genética CRISPR. Este domínio é o mais representativo, visto que os inventores apresentam uma conectividade mais densa, observada pela espessura das ligações. Feng Zhang e Bernd Zetsche apresentam 10 artigos em colaboração sobre CRISPR e suas variações, Cas9 e Cpf1. Eles, bem como Gootenberg são colaboradores do *Broad Institute*, um centro de pesquisa biomédica e genômica localizado em Cambridge, Massachusetts e do Department of Biological Engineering, Massachusetts Institute of Technology. Ambos também aparecem, com menor frequência, acoplados bibliograficamente com Le Cong, Winston X Yan, F Ann Ran e Yinqing Li, cujos temas versam sobre bactéria *Staphylococcus* e Cas9.

Feng Zhang coordena o ZhangLab²⁸, tendo como membro diversos pesquisadores e alunos de iniciação científica e pós-graduação. Entre os integrantes do laboratório, encontra-se Bernd Zetsche, o que nos leva a considerar que esta proximidade é decorrente de uma relação de genealogia acadêmica, uma vez que a atividade de orientação acadêmica é responsável pela transferência e influência de correntes teóricas que formam a identidade do orientando. Nesta perspectiva, as autoras Castanha e Grácio (2015) consideram que os estudos de Genealogia Acadêmica representam relevante abordagem de análise de domínio.

O domínio C3 tem como pesquisadores de destaque Chad Albert Cowan e Mateus Porteus. Ambos não apresentam patentes em comum no corpus do estudo, mas são fundadores científicos do CRISPR Therapeutics, empresa com foco no

²⁸ É possível observar no site do laboratório que a maioria das relações que identificamos das redes, é decorrente de uma relação de genealogia acadêmica. ZhangLab: <https://zlab.bio/alumni>

desenvolvimento de medicamentos baseados em genes a fim de tratar hemoglobinopatias, câncer, diabetes e outras doenças.

Dr. Chad Cowan é Professor Associado da Universidade de Harvard no Departamento de Células Tronco e Biologia Regenerativa e no Hospital Geral de Massachusetts. Matthew Porteus é Professor Associado de Pediatria, Departamento de Pediatria, Divisões de Hematologia/Oncologia e Terapia Genética Humana, na Escola de Medicina de Stanford. Sua pesquisa se concentra no desenvolvimento de terapias baseadas em recombinação homóloga para doenças genéticas. Os autores não apresentam trabalhos em colaboração, mas estão acoplados bibliograficamente por ao menos três diferentes artigos dos pesquisadores, J D Tratschin²⁹, R J Samulski³⁰, O S Platt³¹, que tratam sobre vírus adenoassociados e doença falciforme.

Os inventores acoplados bibliograficamente no cluster C4 fizeram parte, no período que as patentes foram solicitadas, de um centro de pesquisas biotecnológicas, denominado Editas Medicine, localizado em Cambridge, Massachusetts. O cluster C4 tem proximidade com o C2 porque versam sobre mesmos temas e apresentam similitude de autores referenciados.

O cluster C5 tem como domínio o tema células pluripotentes induzidas iPS humanas, além de métodos e técnicas que envolvam modificação direcionada usando RNA. É formado pelos pesquisadores Wojtek Auerbach David M. Valenzuela, membros da Regeneron Pharmaceuticals, Inc. Eles apresentam patentes em comum e ao menos oito publicações científica em colaboração identificadas no Pubmed.

O cluster C6 é formado pelos pesquisadores J. Keith Joung, Benjamin Kleinstiver que trabalham com engenharia genética, edição genômica e sistema CRISPR. Ambos atuam no Massachusetts General Hospital e são docentes no Harvard Medical School. Joung e seus colaboradores desenvolvem tecnologias para engenharia de genoma usando domínios artificiais de ligação de DNA e Kleinstiver procura abordar as limitações das tecnologias de edição do genoma CRISPR.

²⁹ Tratschin, J. D., West, M. H., Sandbank, T., & Carter, B. J. (1984). A human parvovirus, adeno-associated virus, as a eucaryotic vector: transient expression and encapsidation of the procaryotic gene for chloramphenicol acetyltransferase. *Molecular and cellular biology*, 4(10), 2072–2081.

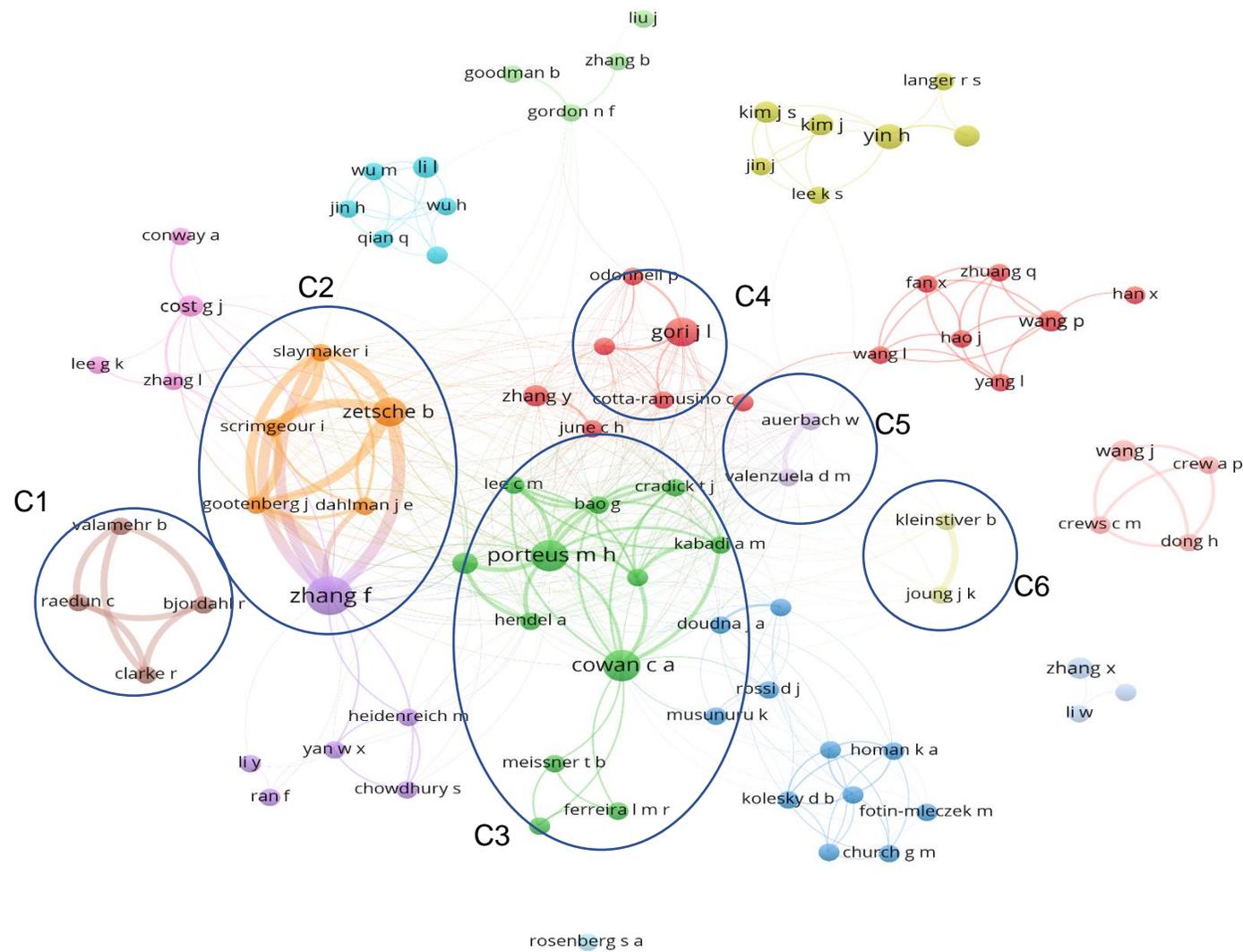
³⁰ Samulski, R. J., Chang, L. S., & Shenk, T. (1989). Helper-free stocks of recombinant adeno-associated viruses: normal integration does not require viral gene expression. *Journal of virology*, 63(9), 3822–3828.

³¹ Platt, O. S., Brambilla, D. J., Rosse, W. F., Milner, P. F., Castro, O., Steinberg, M. H., & Klug, P. P. (1994). Mortality in sickle cell disease. Life expectancy and risk factors for early death. *The New England journal of medicine*, 330(23), 1639–1644.

Nesta rede, foi possível identificar alguns domínios importantes para as tecnologias envolvendo células-tronco. Destaca-se, todavia, a técnica CRISPR e edição genômica como os domínios que predominaram ao longo das análises nas diferentes técnicas metodológicas aplicadas. Além disso, é importante mencionar que Amorin e Café (2016), ao analisar a produção científica de B. Hjørland, salientaram o fato de o conceito de domínio poder abranger um grupo de usuários, uma disciplina e até mesmo uma área do conhecimento, entre outros, e que estes são classificados conforme as necessidades de informação de determinada comunidade discursiva.

Dessa forma, foi possível identificar não apenas o domínio temático, mas também de grupo de pessoas e organizações que sustenta a identidade de citação dos inventores.

Figura 18 – Rede de acoplamento bibliográfico entre os inventores que comportam o domínio do estudo.



Fonte: elaborado pela autora

A Figura 18 nos permitiu identificar as relações existentes entre os inventores e suas proximidades temáticas por meio do acoplamento bibliográfico. No entanto, embora possamos identificar os domínios temáticos e tecnológicos, essas relações nem sempre são claras. Como vimos no cluster C3, é possível afirmar que existe uma relação de proximidade entre orientador x orientando e isso contribui para a definição da identidade de citação do acadêmico. No entanto, quando analisamos os demais clusters, em que identificamos a colaboração e a participação dos inventores no desenvolvimento de tal tecnologia, as razões que os levam a citar a literatura científica permanece não reveladas. O estado da técnica é um dos potenciais motivos para os inventores e os examinadores citarem determinada literatura. Todavia, outros eventuais motivos permanecem ainda não explicitados. Cabe lembrar que o examinador de patentes compara o estado da técnica relacionado ao pedido pendente e, durante este processo, diversas referências podem ser combinadas para mostrar que uma invenção é óbvia e, por isso, não patenteável.

Embora van Raan (2017) considere que a aplicação de técnicas de acoplamento bibliográfico seja importante para melhor identificar os avanços e as tecnologias existentes, o processo para chegar na análise das redes não é simples devido à falta de organização e normalização das informações. Desta forma, considera-se que a razões que levam os inventores e examinadores a citar ainda não estão esclarecidas. Além disso, quando o pesquisador cita, ele constrói a sua identidade de citação, criando um arcabouço teórico importante, mas quando se trata das patentes, devemos levar em consideração que os examinadores também contribuem para a criação desta identidade, visto que são eles que examinam os pedidos de patente e citam diversas literaturas para refutar ou considerar a invenção válida.

6.4 ACOPLAMENTO DE PATENTES SOBRE CÉLULAS-TRONCO

O método elaborado por Kessler (1963) teve como proposta agrupar as publicações científicas, por meio de unidades de acoplamento bibliográfico, ou seja, o acoplamento bibliográfico foi definido pela quantidade de itens de referências em comum entre essas publicações. Por meio desse indicador de similaridade, o autor examina a relação citante – citante. Nesta pesquisa, o método de acoplamento é aplicado para analisar as relações existentes, ou não, entre patentes, em um corpus

de análise constituído pelas 144 patentes registradas entre 2016 e 2020 no campo das células-tronco.

A partir da Figura 19 e da tabela de links de acoplamento (Apêndice C), é possível identificar as relações de acoplamento entre patentes. Grácio e Oliveira (2020) afirmam que a densidade é um indicador destinado a medir a intensidade das conexões das redes. A rede da Figura 19 tem densidade igual a 0,0907 (ou 9,07%) resultante do quociente do número total de ligações (934) nela existente e o número máximo possível de ligações ($10.296 = (144 \times 143) / 2$), evidenciando proximidade de fundamentação tecnológica entre as patentes. Das 144 patentes analisadas, apenas 86 (59,72%) delas apresentaram um acoplamento acima de 3.

A rede de acoplamento formada a partir do VOSviewer, com atração 8 e repulsão -1, gerou 11 clusters que representa um domínio tecnológico dentro da temática células-tronco. Destes clusters, 6 clusters mais evidentes – vermelho composto de 28 patentes, verde com 25 patentes, azul com 20, amarelo com 9 patentes, lilás formado por uma tríade e o laranja formado por 2 patentes. Observa-se que o cluster em vermelho localizado quase que no centro da rede, representado pela patente WO2017023801 é a principal origem de acoplamento entre as demais patentes, uma vez que se acopla a 73 patentes (50,7%) das 144 patentes presentes na rede.

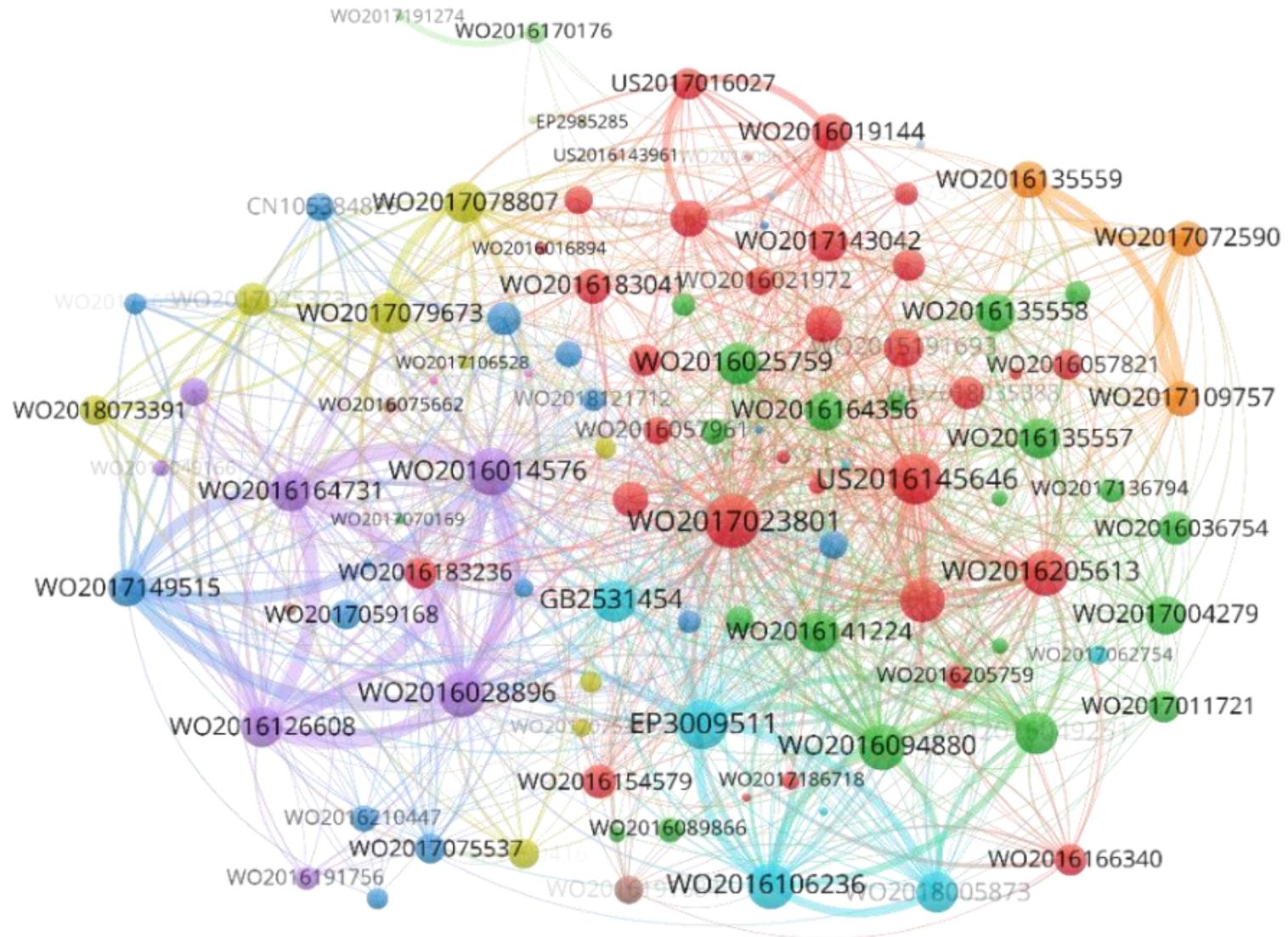
Essa patente tem grande relevância para o processo de intervenção de neoplasias, pois tem como proposta o transplante genômico intracelular e métodos de tratamento. Ela apresenta quatro inventores, sendo o primeiro inventor o professor e pesquisador da Escola de Medicina da *University of Minnesota*, Branden Moriarity, especialista em Ciências Biomoleculares. Conforme eles alegam na patente, apesar dos avanços notáveis na terapêutica do câncer, ainda existem muitos tipos de tumores que são resistentes à quimioterapia, radioterapia ou bioterapia e esta invenção permite reconhecer mutações e apresentar uma resposta imunológica para diversos tipos de câncer.

A patente americana US2016145646 se conecta (acopla) a 61 (42,4%) das 144 patentes analisadas, incluindo a patente acima mencionada. Essa patente trata sobre métodos e composições para modificação genética direcionada usando RNAs, o que indica uma possível similaridade temática entre elas ou um interesse no valor tecnológico que uma tem em relação a outra. Apresenta mais de 10 inventores e tem como cessionária a empresa americana de biotecnologia Regeneron Pharmaceuticals

Inc.³² A empresa investe em parceria com outras instituições envolvidas em pesquisa genética e melhoramento, além de apresentar um centro de pesquisa integrada.

³² Site da empresa: <https://www.regeneron.com/about/collaborations>

Figura 19 – Rede de acoplamento tecnológico entre as patentes pertencentes ao domínio analisado



Fonte: elaborado pela autora.

A terceira patente mais evidente é a EP3009511 que se acopla a 55 (38,2%) patentes, sendo duas delas as patentes WO2017023801 e US2016145646. Essa patente é a que apresenta o maior número de citações recebidas entre as patentes do corpus, o que indica que seja uma invenção com alto potencial tecnológico por tratar de novas enzimas e sistemas CRISPR do pesquisador e inventor Feng Zhang, como mencionado na seção 6.1. As duas patentes mencionadas (WO2017023801 e US2016145646) apresentam similaridade tecnológica com a patente EP3009511, visto que ambas tratam sobre modificações genéticas e métodos genômicos para o tratamento de combate ao câncer.

A quarta patente que tem maior quantidade de outras patentes com similaridade tecnológica é a WO2016014576 ao estar acoplada a outras 49 (34,03%) patentes, entre elas as três patentes já mencionadas. Essa patente tem 15 inventores entre americanos e chineses, sendo a primeira deles a pesquisadora Jennifer Brogdon, que é atualmente Diretora Executiva, chefe de pesquisa em terapia celular, no Departamento de Imuno-Oncologia Exploratória do *Novartis Institutes for BioMedical Research* (NIBR) em Cambridge, MA. Ela é responsável por liderar a equipe que procura desenvolver terapias de células CAR-T (tipo de imunoterapia que envolve a reprogramação das células do sistema imunológico do próprio paciente para combater o câncer) para os tratamentos oncológicos. O cluster de patentes centrado em WO2016014576 é resultante da proximidade temática de tratamento de câncer usando um receptor.

A patente WO2016028896 conectada com outras 48 (33,3%) outras patentes têm como proposta inventiva o uso de células imunes manipuladas para um receptor de antígeno quimérico para tratamentos como leucemias mieloide aguda. Ela está relacionada com todas as patentes até aqui analisadas. Tem como inventora a pesquisadora Jennifer Brogdon apresentando uma invenção colaborativa com pesquisadores da China, Itália e Alemanha, além dos EUA.

Na sequência, estando acoplada com outras 47 (32,64%) patentes encontra-se a patente WO2016094880 que se refere a um sistema, métodos e composições *in vitro*, de distribuição do sistema CRISPR para obter benefícios terapêuticos por meio de edição de genomas de animais. Tem como inventor o médico residente da *Stanford Health Care* - David Benjamin Turitz Cox em colaboração com James Dahlman bioengenheiro e professor da Georgia Institute of Technology e o pesquisador Feng Zhang, pioneiro na pesquisa de CRISPR aplicado aos humanos, tendo como

requerente o centro de pesquisa The Broad Institute Inc., do qual o último inventor é membro.

A sétima patente - WO2016205613 - está acoplada com as 6 primeiras mencionadas e mais 40 outras patentes, totalizando um acoplamento de 46 (31,9%). Feng Zhang, junto com três pesquisadores - Linyi Gao, Bernd Zetsche e Ian Slaymaker – criaram essa invenção sobre a mutação de enzimas CRISPR. Além dela, a patente WO2016106236 – 45 (31,3%) acoplamentos - também trata sobre o sistema CRISPR e tem como inventor o pesquisador Feng Zhang, além dos inventores Patrick Hsu, Jonathan Gootenberg e Aaron Smargon. Essas duas patentes têm como cessionária o centro de pesquisa The Broad Institute.

A nona patente US2016108369 está acoplada as outras 41 (28,5%) patentes desse corpus. Ela trata sobre métodos e composições para gerar ou manter células pluripotentes iPS humanas em cultura. Tem como inventores Junko Kuno, Wojtek Auerbach – especialista em linhagens embrionárias de células-tronco, integrante da empresa cessionária e David M. Valenzuela, vice-presidente da empresa e requerente a empresa Regeneron Pharmaceuticals Inc, que tem como proposta usar camundongos geneticamente humanizados exclusivos para produzir anticorpos totalmente humanos otimizados e anticorpos biespecíficos.

A décima patente com maior número de acoplamento estabelecido entre as patentes da pesquisa é a patente GB2531454, depositada no Reino Unido, com 39 (27,1%) acoplamento. Ela tem como proposta inventiva recuperar a sequência de ácido nucleico útil em medicamentos para o tratamento de doenças, por exemplo, infecção bacteriana ou obesidade. Seu inventor é o Jasper Clube, um advogado de patentes europeu, co-fundador da empresa SNIPRBiome (cessionária da patente) que possui 25 anos de experiência em propriedade intelectual, todos nos setores de biotecnologia e saúde.

As 10 patentes analisadas, com links de acoplamento entre 39 (GB2531454) e 73 (WO2017023801), apresentam uma similaridade tecnológica e temática, uma vez que todas tem como área de conhecimento a Biotecnologia aplicada à Microbiologia e Química, além de todas apresentam no seu roll de classes a classificação internacional de patente C12N, que refere a “microrganismos ou enzimas, suas

composições; propagar, preservar ou manter microrganismos; mutação ou engenharia genética; meios de cultura”³³.

6.5 ANÁLISE DE COCITAÇÃO A PARTIR DOS REFERENTES

A Figura 20 apresenta a rede de cocitação entre as referências que sustentam a base teórica das patentes sobre células-tronco. A rede é formada por 136 referências que receberam no mínimo 10 citações, distribuídas em cinco clusters. A decisão por definir um número mínimo de citações igual a 10 se deu pela visualização e compreensão da rede. Os nós mais evidentes (maiores) sinalizam que as respectivas referências, indicada pelo nome dos autores, tiveram mais menções no portfólio das 144 patentes que compõe este estudo.

Destaca-se que, nesta seção, analisamos a cocitação entre as referências (literatura bibliográfica) presentes no portfólio das 144 patentes que compõe esta tese, ao passo que, na seção 6.6, analisamos a cocitação entre as patentes citadas no portfólio. Como mencionado por Grácio (2020), as análises relacionais de citação podem ser realizadas em nível micro (autores), meso (periódicos, grupos e centros de pesquisa) ou macro (instituições ou países). Nesta tese, optou-se por analisar em nível micro, sem desconsiderar o tipo de material bibliográfico, uma vez que encontramos livros com um número expressivo de citações ao longo da lista de referências.

As referências que comportam o cluster vermelho apresentam uma proximidade temática muito forte no que consiste à edição genômica, engenharia do genoma humano, clonagem molecular, CRISPR-CAs9 e bioinformática. As referências presentes neste cluster podem ser consideradas a imagem de citação estabelecidas pelos inventores e/ou pelos examinadores de patentes, como: US2016145646-A1 sobre métodos e composições para modificação genética direcionada usando RNAs guia pareados, WO2016210447-A1 que trata sobre mascaramento de células t do receptor de antígeno quimérico para ativação específica do tumor, WO2017115268-A1 - composições e métodos para o tratamento de

³³ Informação presente na CIP WIPO

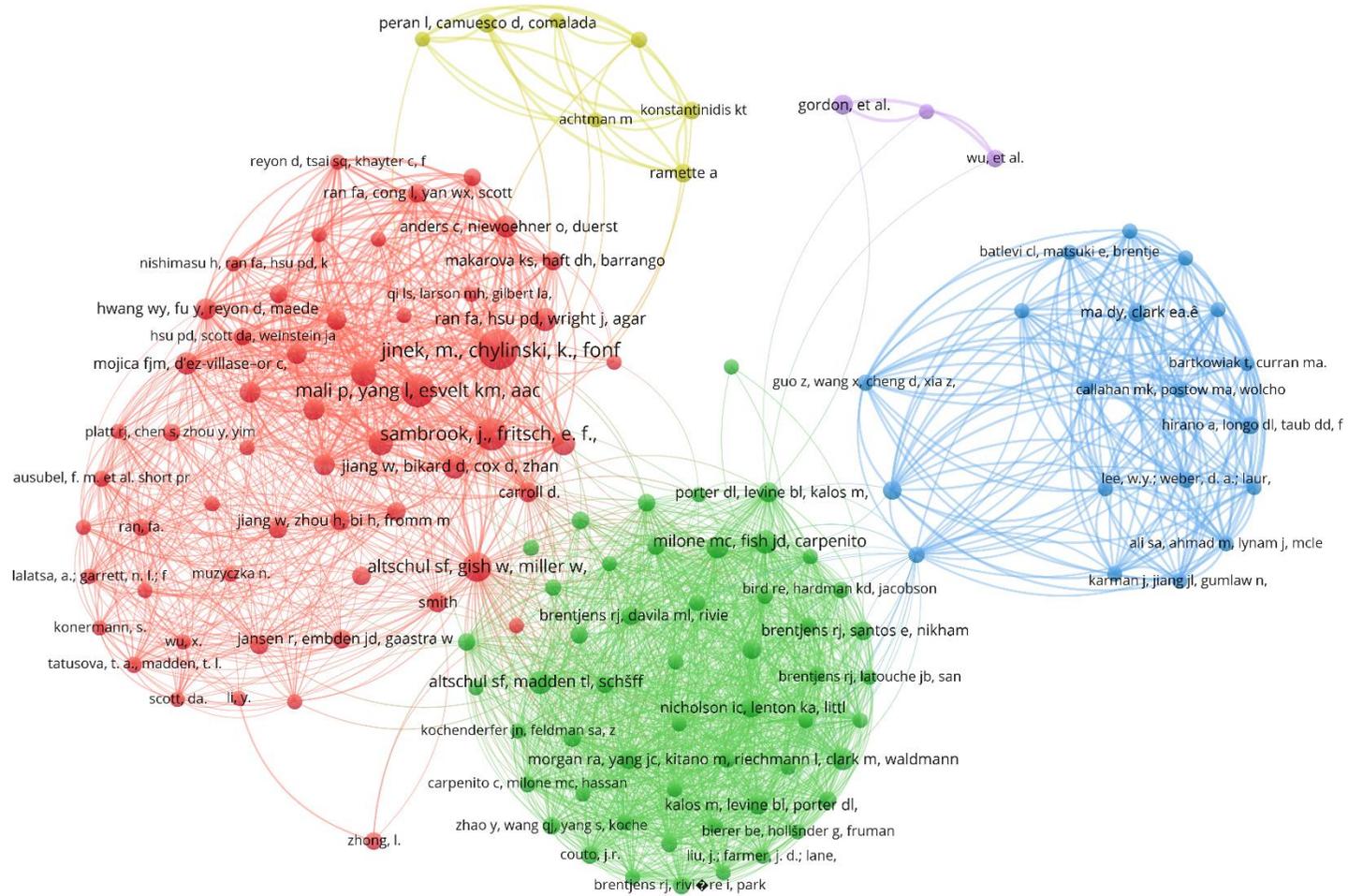
(<https://ipcpub.wipo.int/?notion=scheme&version=20220101&symbol=none&menulang=en&lang=en&viewmode=f&fipopc=no&showdeleted=yes&indexes=no&headings=yes¬es=yes&direction=02n&initial=A&cwid=none&tree=no&searchmode=smart>)

hemoglobinopatias edição genômica e patente WO2016106236-A1 sistema de direcionamento de RNA, esta última patente tem como inventor Feng Zhang. Este cluster indica um domínio voltado para engenharia genética e edição a partir do uso de técnicas acessível como CRISPR CAs9.

As referências do cluster verde são citadas conjuntamente porque tratam de temas como imunoterapia tumoral, células T (também denominadas de linfócitos T), leucemias linfoides, antígenos quiméricos, proteína CD137. A imagem de citação das referências que correspondem a este cluster foram criadas a partir das patentes WO2016028896A1, cuja invenção fornece composições e métodos para o tratamento de doenças associadas à expressão de CD 123; WO2016164731A2 que trata sobre composições e métodos para o tratamento de doenças associadas à expressão de CD19, por exemplo, pela administração de uma célula T recombinante compreendendo o CD19 CAR; WO2017149515A1 que se refere à composição e métodos de células que expressam múltiplas moléculas receptoras de antígenos quiméricos (car) e seus usos, aplicados ao tratamento de doenças ligadas a um antígeno tumoral. Cabe observar que o cluster vermelho e o verde estão relacionados pela referência do autor Altschul e coautores, no National Institutes of Health, responsáveis pelos estudos e elaboração dos algoritmos BLAST que têm como finalidade comparar informações como sequências de aminoácidos de diferentes proteínas ou nucleotídeos de sequências de DNA.

O cluster azul reúne um grupo de referências que abordam temas como linfócitos T, receptores OX40, imunoterapias, terapias CAR-T, oncologia, terapia combinada, CD137, checkpoints (pontos de controle imunológico), entre outros temas que envolvem imunoterapia do câncer. Entre as patentes que fizeram menção às referências deste cluster, encontramos o pedido de patente WO2017109757A1 que apresenta materiais e métodos para tratar um paciente com Esclerose Lateral Amiotrófica (ALS) e/ou Degeneração Lobular Frontaltemporal (FTLD), também o pedido de patente WO2017072590A1 - fornece materiais e métodos para tratar um paciente com Distrofia Muscular de Duchenne (DMD), entre outras associadas à imunoterapias para tratamento de doenças infecciosas, alérgicas, cancerígenas e demais.

Figura 20 - Rede de cocitação da literatura científica referenciada nas patentes



Fonte: elaborado pela autora

Ainda que todas as produções bibliográficas participantes da análise tenham recebido 10 ou mais citações, o cluster lilás é composto por aquelas com menos citações recebidas e trata sobre temas como células modificadas por agentes terapêuticos, receptores CARs – receptores antígenos quiméricos. As referências deste cluster aparecem citadas com a referência do autor Porter D.L. e colaboradores, que tratam sobre o tema dos Linfócitos T modificados por receptor de antígeno quimérico na leucemia linfóide crônica, mencionando Linfócito B e a proteína CD137.

Destaca-se que o cluster amarelo tem uma proximidade temática com o cluster vermelho e cujos autores discorrem sobre diversidade microbiana, bactérias, *Lactobacillus fermentum*, estresse oxidativo intestinal, *Streptococcus*, genomas bacterianos e bioinformática. As proximidades entre os clusters se dá pelo fato de o CRISPR ser uma área do material genético de bactérias que combatem os vírus que as invadem. As bactérias *Streptococcus* apresentam a proteína denominada de Cas9 que combate o DNA viral protegendo a bactéria da infecção, sendo uma das técnicas mais acessível e utilizadas hoje que permite editar qualquer genoma de forma programável.

Sumarizando, identificaram domínios que tratam de técnicas de edição genômicas, mais especificamente CRISPR Cas9, métodos de imunoterapias, além tratamento de distúrbios metabólicos, linfócitos T, sistema de edição de ácido nucleico e nova proteína isolada de *Streptococcus pyogenes*.

Ademais, a rede de cocitação entre as referências mostrou que é possível identificar diferentes domínios e que todos convergem para o domínio engenharia genética voltada a edição de genes com aplicação da técnica CRISPR e suas variações, buscando identificar ferramentas que contribuam para o tratamento das mais diversas doenças, mas dando ênfase para doenças cancerígenas.

6.6 ANÁLISE DE COCITAÇÃO DAS PATENTES SOBRE CÉLULAS-TRONCO

Nesta etapa do trabalho, investigamos a proximidade tecnológica do conjunto de patentes analisadas, conforme reconhecida pela comunidade de inventores e examinadores, à luz da análise de cocitação. Diferente da pesquisa elaborada por Grácio (2020), que realizou análise de estrutura intelectual, este estudo analisou a cocitação entre as patentes citadas no corpus da pesquisa (144 patentes), elucidando

a partir da rede (Figura 21) as principais influências tecnológicas para o desenvolvimento de novas tecnologias.

Grácio (2020) considera que a análise de cocitação permite identificar como uma comunidade discursiva, na presente pesquisa formada por inventores e examinadores das patentes, percebem e reconhecem as contribuições para o avanço do campo e a proximidade entre elas. Breitzman e Moguee (2002) afirmam que no cenário tecnológico a frequência de cocitação entre duas patentes é evidência de que elas estão intimamente relacionadas umas com as outras. Para eles, análises como essas podem ajudar a identificar patentes específicas para licenciamento externo³⁴ e as empresas que possivelmente têm interesse em licenciar. Além disso, contribui para identificar patentes que podem não ter valor significativo para aquela empresa visto que não estão relacionadas com as tecnologias de seu interesse.

Moguee e Kolar (1998) aplicam esse método nas patentes farmacêuticas da Allergan, pois consideram que ele permite analisar a estrutura dos portfólios³⁵ de patentes das empresas e os clusters que se formam, os quais, segundo eles, normalmente, correspondem às áreas de competência tecnológica central. Essa característica é evidenciada pelos resultados obtidos, que revelaram que a Allergan possui três competências tecnológicas centrais em produtos farmacêuticos; que são: compostos com atividade semelhante a retinoides, agentes hipotensores oculares e agentes anti-inflamatórios. Cada uma dessas áreas gerou invenções que estão intimamente relacionadas umas com as outras.

Para analisar esses dados, assim como na rede de acoplamento, consideramos importante calcular a densidade a fim de avaliar a intensidade das conexões e a coesão tecnológica existentes no campo analisado. Grácio e Oliveira (2020) esclarecem que a densidade de uma rede varia entre 0 e 1, em que 0 significa que os atores estão isolados e 1 caracteriza uma rede completa (totalmente conectada). Assim, quanto mais perto de 1 mais densa é a rede. A rede da Figura 18 tem uma densidade baixa de 0,072 (7,25%)³⁶, resultante da presença de 552.651 ligações entre

³⁴ Licenciamento é uma autorização concedida pelo titular da patente a uma ou mais empresas para fabricação e comercialização do produto coberto pela patente. O licenciamento de patente muitas vezes acontece quando o proprietário não possui capacidade produtiva, ou quando ele deseja apenas faturar os royalties pela criação. O licenciamento externo está muito relacionado à falta de interesse da empresa detentora da patente em investir naquele invento e, por sua vez, no potencial tecnológico e valor que ela pode vir a ter para outras empresas que não a titular.

³⁵ É um conjunto de projetos ou produtos desenvolvidos sob gerenciamento de uma organização.

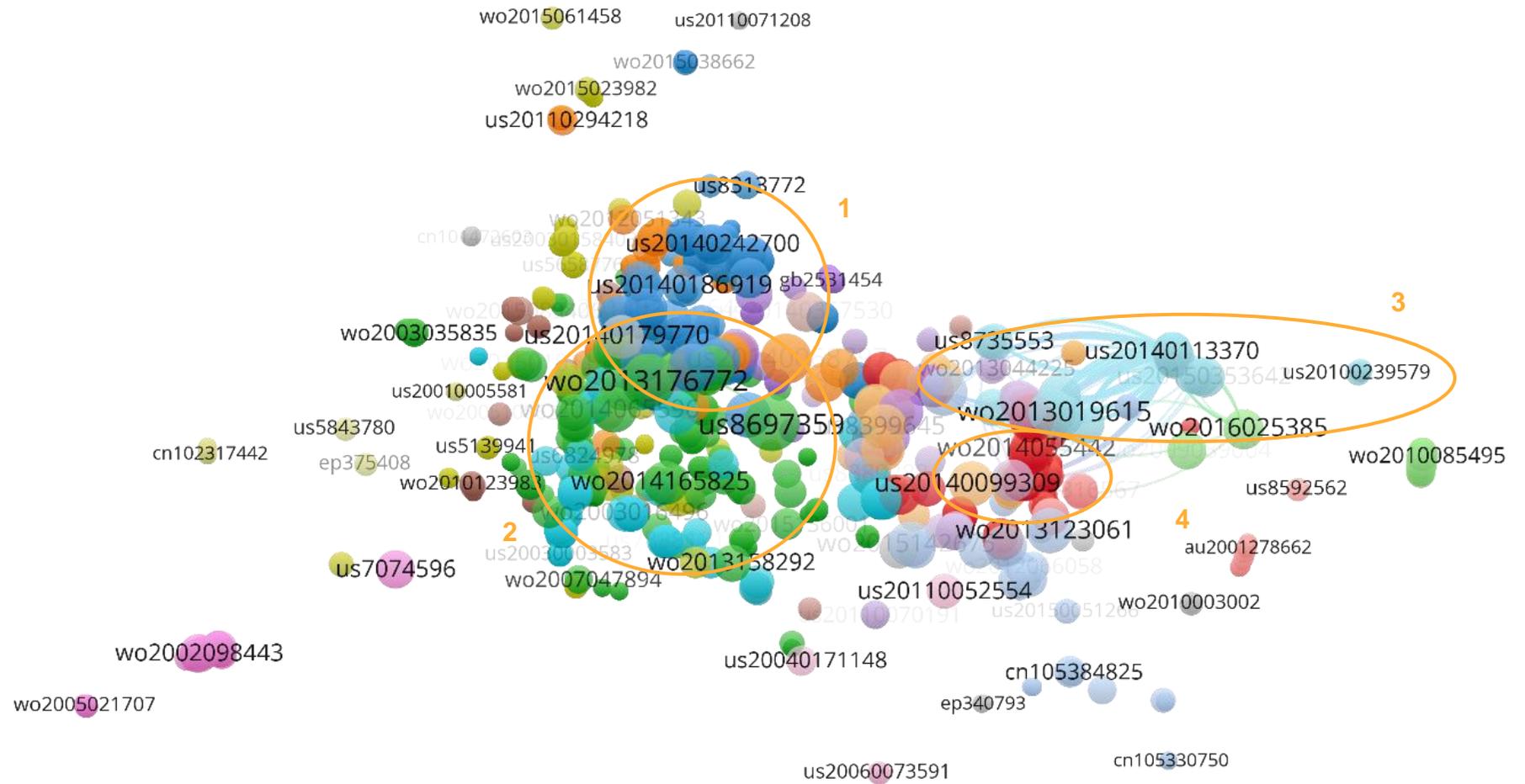
³⁶ Para identificar a lista dos links de cocitação entre as 3902 patentes é possível acessar o arquivo armazenado no drive

as 3902 patentes citadas pelas 144 patentes que compõem o conjunto de documentos da pesquisa. A média de cocitação foi de 286,01, sendo que a patente que mais apresentou cocitação (1825) foi a US8697359 e a que menos apresentou relação de cocitação foi a patente WO2016109668, cocitada uma vez com WO2013123061 e duas vezes com a patente US20140322183.

A rede apresenta 23 clusters, o tamanho dos nós está relacionado à quantidade de citações que cada patente recebeu e não à força de cocitação entre elas. Esse layout é uma característica das redes geradas no software VOSviewer. A US8697359 apresentou maior força de cocitação, sendo cocitada com 1825 outras patentes das outras 3901 tecnologias, e isso se deve ao fato de termos patentes que apresentaram em suas referências mais de 900 citações a outras patentes. Além disso, a patente US8697359 foi cocitada 90 vezes com a patente WO2013176772 e 40 vezes com a patente WO2014093661.

Embora a patente US8697359, cluster verde, tenha apresentado maior força de cocitação entre as demais patentes, a patente US20140113370, cocitada com outras 89 patentes, foi a que apresentou maior grau de proximidade com as demais tecnologias. Ela foi cocitada 144 vezes com a patente WO2005047334 e foi citada conjuntamente 132 vezes com a patente US20150183881.

Figura 21 – Rede de cocitação das patentes citadas pelas 144 patentes do corpus da pesquisa



Fonte: elaborado pela autora.

A fim de elucidar melhor a rede, selecionaram-se quatro clusters que estão em maior evidência e que apresentam uma proximidade maior entre os nós (patentes).

No cluster azul forte (1), as patentes tratam sobre preparações medicinais contendo material genético, além de modificações genéticas de DNA em plantas. Microrganismos ou enzimas; suas composições; propagar, preservar ou manter microrganismos; mutação ou engenharia genética; meios de cultura.

Por sua vez, o cluster verde (2) está voltado às técnicas de engenharia genética em biologia molecular. Abordam mais sobre o sistema CRISP, que permite a edição de genes defeituosos, sendo uma técnica moderna de edição genética de diversos organismos, incluindo plantas. A sobreposição observada entre os clusters azul e verde (1 e 2) evidencia a proximidade temática e tecnológica das patentes desses clusters.

Wang e Duan (2011) compreendem que se as patentes A e B são cocitadas, elas têm correlação nas tecnologias e a correlação aumenta com o aumento da frequência cocitada. Isso fica evidente também no cluster azul claro (3), cujas patentes apresentaram maior frequência de cocitação US20140113370 e WO2005047334 US20150183881, como mencionado anteriormente. Além disso, as patentes que compõe esse cluster tratam sobre anticorpos, para uso terapêutico e diagnósticos, além de tratar sobre proteína de fusão³⁷ ou proteína híbrida, bem como de composições e Métodos Relacionados para o Tratamento de Doenças e Distúrbios Imunológicos utilizando proteínas de fusão.

Por sua vez, o cluster em vermelho (4) trata sobre composições e métodos para o tratamento de câncer em humano, método imunoterapêutico adotivo de tratamento de um paciente com um antígeno específico de tumor portador de câncer. Métodos de tratamento de câncer abordam grupos de Células T que desempenham um papel importante no combate a infecções e células tumorais.

Mogee e Kolar (1998; 1999) empregaram informações de cocitação de patentes para identificar as tecnologias centrais da Allergan Inc e Eli Lilly & CO. Lai e Wu (2005) estabeleceram um novo sistema de classificação de patentes usando informações de cocitação de patentes. Wu, Chen e Lee (2010) usaram a cocitação de patentes para revelar as principais tecnologias da TSMC Co. Dessa forma, Wang e Duan (2011)

³⁷ É uma proteína artificial obtida a partir de um gene de fusão, o qual é criado ao juntar partes de dois genes diferentes. Fonte: National Cancer Institute (<https://www.cancer.gov/publications/dictionaries/cancer-terms/def/fusion-protein?redirect=true>)

usaram o método de cocitação de patentes para identificar a tecnologia principal da indústria de veículos elétricos.

A rede de cocitação das patentes citadas mostrou as principais influências tecnológicas para o desenvolvimento de novas tecnologias no cenário de células-tronco, em que as temáticas com maior influência (citação) estão relacionadas ao melhoramento genético tanto para humanos como plantas, buscando apresentar métodos de tratamento para doenças autoimunes e neoplasias

7 CONSIDERAÇÕES FINAIS

Para avaliar a viabilidade da metodologia em documentos de patentes, escolhemos aplicá-la às tecnologias voltadas às células-tronco, visto que é uma temática importante nas diversas áreas que compõe a Ciências da Saúde, tanto para o desenvolvimento de tratamentos oncológicos, como para qualidade de vida e estética. No cenário da Inovação e Tecnologia, indústrias que apresentam capacidade tecnológica avançada, detêm maior vantagem competitiva no mercado, promovendo benefícios econômicos para si e para o país. Van Raan (2017) considera que a análise de cocitação e de acoplamento bibliográfico fornece uma maneira confiável, eficaz e muito menos demorada de descobrir importantes fluxos de conhecimento entre ciência e tecnologia, identificar as publicações e patentes que desempenham um papel fundamental nesses fluxos e encontrar os primeiros sinais de temas tecnológicos emergentes.

Esta pesquisa desenvolveu-se a partir da indagação alusiva à contribuição das análises uni-variada e relacional de citação em documentos de patente como método para identificação de domínios tecnológicos relacionados às patentes em células tronco, representados nas análises pelas redes de acoplamento (11 domínios/clusters) e cocitação (23 domínios/clusters). Para isso, trabalhou-se com um conjunto de 144 patentes que tratam do assunto células-tronco, bem como as referências às patentes e aos documentos não patentes, citadas tanto pelo inventor como pelo examinador. Usou-se como recorte as patentes mais citadas, indexadas na base de dados Derwent Innovation Index no período de 2016 a 2020.

Neste sentido, analisando o corpus da pesquisa, composto por 144 patentes, observou-se que a patente EP3009511 foi a patente mais citada, com 297 citações. Ela é apresenta na sua lista de referências, elaborada tanto pelo inventor como pelo examinador, 681 patentes citadas e 1104 referências de outros tipos de documentos indicando um extenso levantamento de publicações tanto científicas como tecnológicas sobre o assunto (Apêndice B). Ela trata sobre novas enzimas do sistema CRISPR - sequências de ácidos nucleicos repetitivas, que funcionam como um sistema de defesa antiviral. A invenção fornece sistemas, métodos e composições para direcionar ácidos nucleicos, sendo a mais modernas técnicas de edição genética aplicada em diversos organismos, entre eles plantas, contribuindo para o melhoramento genético na agricultura, caso da patente mencionada.

Embora o sistema CRISPR não tenha sido cunhado pelo bioengenheiro sino-americano Feng Zhang³⁸, especialista em optogenética³⁹, foi ele o primeiro pesquisador a patentear a técnica voltada a seres humanos.

Entre as 10 patentes citantes que receberam maior número de citações (Apêndice B), outras duas (WO2016205613, WO2016106236) também têm como inventor o pesquisador Feng Zhang e tratam sobre a mesma temática. Além dessas, outra que trata sobre CRISPR é a patente (US2016145646) – 59 citações – cuja invenção foi desenvolvida pelo doutor em bioquímica David Frendewey, que atua como diretor Executivo de Descoberta e Triagem da Regeneron Pharmaceuticals. Indicando um forte interesse tecnológico sobre essa técnica.

Ainda, foi possível observar que entre as 4915 patentes citadas no corpus da pesquisa, a patente mais citada entre as citadas (US8697359, com 35 citações) fornece sistemas, métodos e composições para alterar a expressão de sequências de genes alvo e produtos de genes relacionados e tem como único inventor o pesquisador Feng Zhang. Dessa forma, o que se observa entre as 10 primeiras patentes que compõem as 4915 patentes citadas, as 5 primeiras estão mais voltadas à temática CRISPR, enquanto as outras 5 patentes mencionadas acima tratam de receptores que buscam aumentar a resposta imunológica e contribuir no tratamento de doenças neoplásicas.

Sobre a análise do referencial científico, foi possível observar que mesmo nas patentes há um número muito expressivo de autores pouco citados e um número pequeno de autores muito citados. No entanto, considerando que o autor (referência) que recebeu maior citação foi Jinek M e coautores, com 74 citações, podemos considerar que as citações à literatura científica em documentos de patente são importantes, mas pouco representativas quantitativamente, uma vez que neste caso a referência mais citada corresponde a 0,51% das citações.

No que concerne a análise de acoplamento bibliográfico, observa-se que a rede composta pelo cluster em vermelho localizado no centro, representado pela patente WO2017023801 é a principal origem de acoplamento entre as demais patentes, uma vez que se acopla a 73 patentes (50,7%) das 144 patentes presentes na rede, sendo relevante para o processo de intervenção de neoplasias. Apresenta quatro inventores,

³⁸ Informações fornecidas no site da Beam Therapeutics (<https://beamtx.com/team/feng-zhang/>)

³⁹ Combinação de métodos genéticos e ópticos no controle de eventos específicos com precisão temporal em células alvo fe um sistema biológico intacto e funcional.

sendo o primeiro inventor o professor e pesquisador da Escola de Medicina da *University of Minnesota*, Branden Moriarity, especialista em Ciências Biomoleculares. Conforme eles alegam na patente, apesar dos avanços notáveis na terapêutica do câncer, ainda existem muitos tipos de tumores que são resistentes à quimioterapia, radioterapia ou bioterapia e esta invenção permite reconhecer mutações e apresentar uma resposta imunológica para diversos tipos de câncer.

A patente americana US2016145646 apresenta proximidade de acoplamento com outras 61(42,4%) patentes das 144 analisadas, incluindo a patente acima mencionada. Ela trata sobre métodos e composições para modificação genética direcionada usando RNAs, o que indica uma possível similaridade temática entre elas ou um interesse no valor tecnológico que uma tem em relação a outra.

As 10 patentes analisadas, com links de acoplamento entre 39 (GB2531454) e 73 (WO2017023801), apresentam uma similaridade tecnológica e temática, uma vez que todas tem como área de conhecimento a Biotecnologia aplicada a Microbiologia e Química, além de todas apresentam no seu roll de classes a classificação internacional de patente C12N, que refere a “microrganismos ou enzimas, suas composições; propagar, preservar ou manter microrganismos; mutação ou engenharia genética; meios de cultura.

No acoplamento bibliográfico de inventores, utilizando como unidade de análise as referências à literatura não patente (literatura científica) – campo CR da DII, observaram-se diversos domínios, como grupo de indivíduos que colaboram entre si e atuam em um mesmo laboratório de pesquisa ou em uma mesma indústria do segmento de biotecnologia. Destacamos os domínios representados pelos clusters: C1, cujos colaboradores atuam na indústria de biofarmacologia e apresentam patentes juntos; C2 em que identificamos a presença da temática edição genômica e CRISPR bem presentes e a participação do pesquisador Feng Zhang, líder de um laboratório de pesquisa, com artigos publicados, que solicita patente com seus colaboradores, indicando uma relação de genealogia acadêmica entre eles, o que afeta diretamente na fundamentação teórica dos seus coautores/coinventores. Além disso, identificamos um domínio (C3) cujos inventores não apresentam nenhuma colaboração científica em comum, mas acoplados bibliograficamente pelas similaridades de referências, visto que ambos pesquisam o mesmo tema e fundaram, juntos, a empresa CRISPR Therapeutics.

Na seção 6.5, a análise de cocitação a partir dos referentes evidenciou cinco clusters bem definidos e que dialogam entre si por meio de uma referência ou outra. Na rede (Figura 20), identificou-se que as referências cocitadas discorriam sobre temas similares, mas com alguma técnica ou abordagem diferente. Destacamos o cluster vermelho, cujas referências cocitadas tratam sobre edição genômica, clonagem molecular, CRISPR-CAs9 e bioinformática. Esse cluster está muito relacionado às patentes que envolvem técnicas e métodos de edição genômica e aplicação do CRISPR. As referências do cluster verde são citadas conjuntamente por tratarem de temas como imunoterapia tumoral e células T. Está próxima ao cluster azul, apresentando um referente em comum, porque trata dos mesmos temas, mas sob perspectivas e técnicas diferentes.

No que tange a análise de cocitação das patentes citadas pelas 144, desenvolveu-se uma rede com 23 clusters. A US8697359 apresentou maior força de cocitação, sendo cocitada com 1825 outras patentes das outras 3901 tecnologias, e isso se deve ao fato de termos patentes que apresentaram em suas referências mais de 900 citações a outras patentes. Além disso, a patente US8697359 foi cocitada 90 vezes com a patente WO2013176772 e 40 vezes com a patente WO2014093661.

Embora a patente US8697359, cluster verde, tenha apresentado maior força de cocitação entre as demais patentes, a patente US20140113370, cocitada com outras 89 patentes, foi a que apresentou maior grau de proximidade com as demais tecnologias. Ela foi cocitada 144 vezes com a patente WO2005047334 e foi citada conjuntamente 132 vezes com a patente US20150183881.

A fim de elucidar melhor a rede, selecionaram-se 4 clusters que apresentam uma proximidade maior entre os nós. Entre os clusters, o que está em maior evidência é o azul forte, que tratam sobre preparações medicinais contendo material genético, além de modificações genéticas de DNA em plantas. Microrganismos ou enzimas; suas composições; propagar, preservar ou manter microrganismos; mutação ou engenharia genética; meios de cultura. O cluster verde está voltado às técnicas de engenharia genética em biologia molecular. Abordam mais sobre o sistema CRISPR, que permite a edição de genes defeituosos, sendo uma técnica moderna de edição genética de diversos organismos, incluindo plantas. Essa sobreposição entre os clusters azul e verde mostrou que há proximidade temática e tecnológicas das patentes.

Já o cluster azul claro, cujas patentes apresentaram maior frequência de cocitação US20140113370 e WO2005047334 US20150183881, como mencionado anteriormente. Além disso, as patentes que compõem este cluster tratam sobre anticorpos, para uso terapêutico e diagnósticos, além de tratar sobre proteína de fusão ou proteína híbrida, bem como, de composições e métodos relacionados ao tratamento de doenças e distúrbios imunológicos, utilizando proteínas de fusão.

Por sua vez, o cluster em vermelho voltados para métodos e tratamentos específicos de tumor portador de câncer. Métodos de tratamento de câncer. Abordam grupos de Células T que desempenham um papel importante no combate a infecções e células tumorais.

A partir das análises das redes de cocitação, identificou-se que os 23 clusters, representam domínios epistêmicos que configuram as correntes tecnológicas em que atuam os inventores participantes dos domínios tecnológicos em células-tronco evidenciados na Figura 21.

O estudo mostrou que os métodos bibliométricos uni-variados e relacionais de citação, são aplicáveis para identificação de domínios tecnológicos, ratificando a hipótese apresentada no início do estudo. No entanto, observa-se que para melhor caracterizar o domínio tecnológico, deve-se aplicar conjuntamente aos estudos bibliométricos, o uso e/ou desenvolvimento de tesouro e análise qualitativa, buscando o conhecimento especializado. Ainda, considerando que pesquisadores como Pimenta (2017) e Ravaschio, Faria e Quoniam (2010) veem as patentes como fontes de informação e observando as diversas limitações apresentadas no estudo, cabe ponderar sobre a importância de aplicar estudos envolvendo sistema de organização do conhecimento, recuperação da informação, terminologia e demais.

No decorrer do desenvolvimento da pesquisa, observa-se uma desordem informacional relativa às referências citadas no campo CR, campo de referências aos documentos não-patente. Esta ausência de normalização limita as análises e a busca pelo documento citado, tornando o processo de limpeza dos dados moroso e passível de erros, visto que é feito manualmente. A pesquisa foi desenvolvida utilizando como instrumento de análise as patentes, com aplicação de técnicas bibliométricas como cocitação e acoplamento bibliográfico. Essas técnicas são consideradas importantes e viáveis por van Raan (2017) para análise de estudos de patente. No entanto, considerando que a pesquisa é realizada na área de Ciência da Informação, compreende-se que ao observar as problemáticas apontadas no estudo, pondera-se

que os escritórios de patentes são um nicho de mercado de trabalho pouco considerados ou desconhecidos pelos profissionais de CI. A partir disso, a pesquisa também propõe que estudos como este sejam aplicados em diferentes bases de dados, a fim de identificar se estas incongruências se repetem e, com isso, propor um processo de normalização e campos obrigatórios que devem ser alimentados.

Finalizando, sugerem-se pesquisas futuras que aliem a modelagem de tópicos à aplicação desta metodologia, para melhor compreender os domínios que contempla as áreas de conhecimento do pesquisador. A modelagem de tópicos é uma estratégia para identificar e organizar grandes bancos de informações, utilizando algoritmos e métodos estatísticos que identificam temas e suas relações. Entre os algoritmos capazes de serem utilizados para extração de tópicos do grande volume de dados existe o *Latent Semantic Indexing* (LSI) e *Latent Dirichlet Allocation* (LDA). O último é o modelo mais utilizado para organizar grandes volumes de documentos, desta forma, observando a complexidade dos documentos de patentes, considera-se um método apropriado para análises futuras.

REFERÊNCIAS

- ALMEIDA, J. F. V. R.; DIAS, G. A. Estado da arte sobre análise de domínio no campo da Ciência da Informação brasileira. **Brazilian Journal of Information Science: research trends**, [s. l.], v. 13, n. 3, p. 26–45, 2019. DOI: 10.36311/1981-1640.2019.v13n3.04.p26. Disponível em: <https://revistas.marilia.unesp.br/index.php/bjis/article/view/8610>. Acesso em: 20 maio. 2022.
- AMORIM, I. S.; VIANNA, W. B.; BASILIO MEDEIROS, M. B. Aspectos epistemológicos em estudos sobre análise de domínio sob as lentes do racionalismo e do empirismo. **Logeion: Filosofia da Informação**, [s. l.], v. 5, n. 1, p. 35-47, 2018. DOI: 10.21728/logeion.2018v5n1.p35-47. Disponível em: <https://revista.ibict.br/fiinf/article/view/4229>. Acesso em: 20 maio 2022.
- ÁVILA-ROBINSON, A.; SENGOKU, S. Tracing the knowledge-building dynamics in new stem cell technologies through techno-scientific networks. **Scientometrics**, Dordrecht, v. 112, n. 3, p. 1691-1720, 2017.
- BASSECOULARD, E.; ZITT, M. Patents and publications: the lexical connection. *In*: MOED, H. F.; GLÄNZEL, W.; SCHMOCH, U. (ed.). **Handbook of quantitative science and technology research**. Netherlands: Kluwer Academic Press, 2004.
- BAYER, A. E.; SMART, J. C.; McCLAUGHLIN, G. W. Mapping intellectual structure of a scientific subfield through author cocitations. **Journal of the American Society for Information Science**, [s. l.], v. 41, n. 6, p. 444-452, 1990.
- BEGHTOL, C. Domain analysis, literary warrant, and consensus: the case of fiction. **Journal of the American Society for Information Science**, [s. l.], v.46, n.1, jan.1995.
- BERTIN, P. R. B.; VISOLI, M. C.; DRUCKER, D. P. A gestão de dados de pesquisa no contexto da e-science: benefícios, desafios e oportunidades para organizações de P&D. **Ponto de Acesso**, Salvador, v. 11, n. 2, p. 34-48, ago. 2017. Disponível em: <https://ainfo.cnptia.embrapa.br/digital/bitstream/item/168396/1/A-gestao-de-dados-de-pesquisa....pdf>. Acesso em: 04 jun. 2018.
- BOCHI, F.; GABRIEL JÚNIOR, R. F.; MOURA, A. M. M. Patentes nos estudos métricos da informação. *In*: GRÁCIO, M. C. C. *et al.* (org.). **Tópicos da bibliometria para bibliotecas universitárias**. Marília: Oficina Universitária, 2020.
- BORGMAN, C. Research data: who will share what, with whom, when, and why? *In*: CHINA--NORTH AMERICAN LIBRARY CONFERENCE, 5., 2010, Beijing. **Anais [...]** Beijing: [s. n.], 2010. Disponível em: <https://works.bepress.com/borgman/238/>. Acesso em: 04 jun. 2018.
- BORNMANN, L.; DANIEL, H. D. What do citation counts measure? A review of studies on citing behavior. **Journal of Documentation**, [s. l.], v. 64, n. 1, p. 537-607, 2007.

BRASIL. **Lei nº 9.279, de 14 de maio de 1996.** Regula direitos e obrigações relativos à propriedade industrial. Brasília, DF: Presidência da República, 1996. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/l9279.htm. Acesso em: 20 maio 2022.

BRASIL. Ministério da Economia. **Manual de Básico para Proteção por Patentes Invenções, Modelos de Utilidade e Certificados de Adição.** Brasília, DF: INPI, 2021.

BRASIL. Ministério do Desenvolvimento, Indústria e Comércio Exterior. **Diretrizes de exames de pedidos de patentes na área de biotecnologia.** Brasília, DF: INPI, 2015.

BREITZMAN, A. F.; MOGEE, M. E. The many applications of patent analysis. **Journal of Information Science**, [s. l.], v. 28, p. 187-205, 2002.

BUFREM, L.; PRATES, Y. O saber científico registrado e as práticas de mensuração da informação. **Ciência da Informação**, Belo Horizonte, v. 34, n. 2, p. 9-25, 2005.

CAETANO, G. C. G. *et al.* Técnica CRISPR-CAS9 e sua utilização na área laboratorial. **Brazilian Journal of Surgery and Clinical Research**, Minas Gerais, v. 25, n. 2, p. 96-99, 2018.

CALLON, M.; COURTIAL, J-P.; PENAN, H. **Cienciometría: la medición de la actividad científica - de la bibliometría a la vigilancia tecnológica.** Asturias: Ediciones Trea, S.L., 1995.

CAMARGO, J. R. F. **Aproveitamento da informação tecnológica em pesquisas acadêmicas:** análise de citações de patentes em teses e dissertações da área de engenharia. 2011. 181 f. Dissertação (Mestrado em Multidisciplinar) – Universidade Federal de São Carlos, São Carlos, 2011.

CAMP, L. S. **A história secreta e curiosa das grandes invenções.** Rio de Janeiro: Lidoador, 1961.

CAMPOS, A. C.; DENIG, E. A. Propriedade intelectual: uma análise a partir da evolução das patentes no Brasil. **Revista Faz Ciência**, Francisco Beltrão, v. 13, n. 18, p. 97-120, jul./dez. 2011. DOI: doi.org/10.48075/rfc.v13i18.7977. Disponível em: <https://e-revista.unioeste.br/index.php/fazciencia/article/view/7977/6700>. Acesso em: 31 maio 2022.

CANALLI, W. M.; SILVA, R. P. Uma breve história das patentes: analogias entre ciência x tecnologia e trabalho intelectual x trabalho operacional. *In*: SCIENTIARUM HISTÓRIA, 4., 2011, Rio de Janeiro. **Anais [...]**. Rio de Janeiro: UFRJ, 2011. v. único. p. 742-748.

CANTOS-MATEOS, G. *et al.* Stem cell research: bibliometric analysis of main research areas through KeyWords Plus. **Aslib Proceedings**, [s. l.], v. 64, n. 6, p. 561-590, 2012.

CASTANHA, R.; ROSAS, F.; GRÁCIO, M. C. C. The Complementarity of Hjørland's and Tennis's proposals to domain analysis under bibliometrics. **Advances in Knowledge Organization**, [s. l.], v. 15, p. 219-226, 2016.

CASTANHA, R.G. The-Bibliographic-Coupler v1.1.1. **Zenodo**, [s. l.], 25 jan. 2022. Disponível em: <https://zenodo.org/record/5903774#.YpkiAcXMLcc>. Acesso em: 31 maio 2022.

CASTANHA, R. C. G.; GRÁCIO, M. C. C. **Estudos de genealogia acadêmica como abordagem para análise de domínio**. , p. 108-116, . Disponível em: <http://hdl.handle.net/20.500.11959/brapci/135205>. Acesso em: 27 mar. 2023.

CASTANHAS, R. G.; BUFREM, L. S.; BOCHI, F. Estudos relacionais de citação: cocitação, acoplamento bibliográfico e genealogia científica. *In*: GRÁCIO, M.C.C. *et al.* (org.). **Tópicos da bibliometria para bibliotecas universitárias**. Marília: Oficina Universitária, 2020.

CATIVELLI, A. S. **Indicadores métricos de valor de patentes: construção de um Índice de Valor utilizando as patentes verdes brasileiras**. 2020. Tese (Doutorado em Ciência da Informação) – Universidade Federal de Santa Catarina, Florianópolis, 2020.

CHAVES, G. C. *et al.* A evolução do sistema internacional de propriedade intelectual: proteção patentária para o setor farmacêutico e acesso a medicamentos. **Caderno de Saúde Pública**, Rio de Janeiro, v. 23, n. 2, p. 257-267, fev. 2007.

CHEN, H.; ZHANG, G.; LU, J.; ZHU, D. A fuzzy approach for measuring development of topics in patents using Latent Dirichlet Allocation. *In*: IEEE INTERNATIONAL CONFERENCE ON FUZZY SYSTEMS (FUZZ-IEEE), 2015, Istanbul. **Anais [...]**. Istanbul: IEEE, 2015. p. 1-7. DOI: 10.1109/FUZZ-IEEE.2015.7337980. Disponível em: <https://ieeexplore.ieee.org/document/7337980>. Acesso em: 31 maio 2022.

COELHO, F. U. **Manual de direito comercial**. 9. ed. São Paulo: Saraiva, 1997.

COLE, F. J.; EALES, N. B. The history of comparative anatomy: part 1 - a statistical analysis of the literature. **Science Progress**, [Londres], v. 11, n. 44, p. 578-596, abr. 1917.

CONSONI, L. A. E. A. **Produção tecnológica em biodiesel: análise das características dos depósitos de patentes indexadas na Derwent Innovations Index entre 1983 e 2015**. 2017. 193 f. Dissertação (Mestrado em Comunicação e Informação) – Programa de Pós-Graduação em Comunicação e Informação, Faculdade de Biblioteconomia e Comunicação, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2017.

COOK, Trevor. Stem Cell Patenting in the European Union. **Journal of Intellectual Property Rights**, [s. l.], v. 17, p. 73-75, jan. 2012.

CORRÊA, M. V. O admirável projeto genoma humano. **Physis: Revista de Saúde Coletiva**, Rio de Janeiro, v. 12, n. 2, p. 277-299, 2002.

CRONIN, B.; SHAW, D. Identity-creators and image-makers: using citation analysis and thick description to put authors in their place. **Scientometrics**, Dordrecht, v. 54, n. 1, p. 31-49, 2002.

CWTS. Welcome to VOSviewer. Leiden: CWTS, 2021. Disponível em: <https://www.vosviewer.com/>. Acesso em: 20 jun. 2022.

DINIZ, D.; AVELINO, D. Cenário internacional da pesquisa em células-tronco embrionárias. **Revista de saúde pública**, São Paulo, v. 43, n. 3, p. 541-547, 2009.

DONALD, K. E.; KABIR, K.; DONALD, W. A. Tips for reading patents: a concise introduction for scientists. **Expert opinion on therapeutic patents**, [s. l.], v. 28, n. 4, p. 277-280, 2018. DOI: doi.org/10.1080/13543776.2018.1438409. Disponível em: <https://www.tandfonline.com/doi/full/10.1080/13543776.2018.1438409>. Acesso em: 31 maio 2022.

EGGER, Peter H.; SELIGER, Florian; WOERTER, Martin. On the distribution of patent citations and its fundamentals. **Economics Letters**, [s. l.], v. 147, p. 72-77, 2016.

EOM H., CHOI S., CHOI S. O. Marketable value estimation of patents using ensemble learning methodology: Focusing on U.S. patents for the electricity sector. **PLoS ONE**, São Francisco, v. 16, n. 9, 2021.

ÉRDI, P.; MAKOVI, K.; SOMOGYVÁRI, Z. *et al.* Prediction of emerging technologies based on analysis of the US patent citation network. **Scientometrics**, Dordrecht, n. 95, p. 225-242, 2013.

EVANGELISTA, I. V.; GRÁCIO, M. C. C.; GUIMARÃES, J. A. C. The Concepts of Domain, Discourse Community and Epistemic Community: affinities and specificities. **Brazilian Journal of Information Science: research trends**, [s. l.], v. 16, p. e02138, 2022. DOI: [10.36311/1981-1640.2022.v16.e02138](https://doi.org/10.36311/1981-1640.2022.v16.e02138). Disponível em: <https://revistas.marilia.unesp.br/index.php/bjis/article/view/12591>. Acesso em: 20 maio 2022.

FEDERMAN, S. R. **Patentes: Desvendando seus Mistérios**. Rio de Janeiro: Qualitymark, 2006.

FREITAS, R. E.; ZUCOLOTO, G. F. Propriedade Intelectual e Aspectos Regulatórios Em Biotecnologia: Estados Unidos. *In*: ZUCOLOTO, G. F.; FREITAS, R. E. (org.) **Propriedade Intelectual e aspectos regulatórios em Biotecnologia**. Rio de Janeiro: IPEA, 2013.

GABRIEL JUNIOR, R. F. **Geração de indicadores de produção e citação científica em revistas de Ciência da Informação: estudo aplicado à base de dados BRAPCI**. 2014. 140 f. Dissertação (Mestrado Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista Julio de Mesquita Filho, 2014.

GABRIEL JUNIOR, R. F.; LAIPELT, R. C. Thesa: ferramenta para construção de tesouro semântico aplicado interoperável. **Revista P2P & Inovação**, Rio de Janeiro, v. 3, n. 2, p.124-145, 2017.

GARFIELD, E. Citation analysis as a tool in journal evaluation: journals can be ranked by frequency and impact of citations for science policy studies. **Science**, [s. l.], v. 178. p. 471-479, 1972.

GILBERT, G.N. Referencing as Persuasion. *Social Studies of Science*, Vol. 7, No. 1 (Feb., 1977), pp. 113-122 (10 pages). <https://www.jstor.org/stable/284636>

GLÄNZEL, W.; MEYER, M. Patents cited in the scientific literature: an exploratory study of 'reverse' citation relations. **Scientometrics**, Dordrecht, v. 58, n. 2, p. 415-28, 2003.

GRÁCIO, M. C. C. Acoplamento bibliográfico e análise de cocitação: revisão teórico-conceitual. **Encontros Bibli**: revista eletrônica de biblioteconomia e ciência da informação, Florianópolis, v. 21, n. 47, p. 82-99, set./dez. 2016

GRÁCIO, M. C. C. **Análises relacionais de citação para a identificação de domínios científicos**: uma aplicação no campo dos Estudos Métricos da Informação no Brasil. 2018. Tese (Livre Docência em Estudos Métricos da Informação) – Faculdade de Filosofia e Ciências - Universidade Estadual de São Paulo, Marília, 2018.

GRÁCIO, M. C. C. **Análises relacionais de citação para a identificação de domínios científicos**: uma aplicação no campo dos Estudos Métricos da Informação no Brasil. São Paulo: Universidade Estadual de São Paulo: Oficina Universitária, 2020.

GRÁCIO, M. C. C.; OLIVEIRA, E. F. T. Análise de cocitação de autores: um estudo teórico-metodológico dos indicadores de proximidade, aplicados ao GT7 da ANCIB. **Liinc em Revista**, Rio de Janeiro, v. 9, n. 1, p. 196-213, maio 2013. Disponível em: <http://revista.ibict.br/liinc/index.php/liinc/article/view/527/401>. Acesso em: 30 jun. 2021.

GRÁCIO, M. C. C.; OLIVEIRA, E. F. T. Análise de redes sociais para visualização do comportamento científico. *In*: GRÁCIO, M.C.C. *et al.* (org.). **Tópicos da bibliometria para bibliotecas universitárias**. Marília: Oficina Universitária, 2020.

GRÁCIO, M. C. C.; OLIVEIRA, E. F. T. Estudos de Análise de Cocitação de Autores: uma abordagem teórico-metodológica para a compreensão de um domínio. **Tendências da Pesquisa Brasileira em Ciência da Informação**, Rio de Janeiro, v. 7, n. 1, p. 1-22, 2014.

GUIMARAES, J. A. C. Análise de domínio como perspectiva metodológica em organização da informação. **Ciência da Informação**, Brasília, DF, v. 43, n. 1, p. 13-21, 2014.

GUIMARÃES, J. A. C. *et al.* A dimensão teórica da análise de domínio na produção científica brasileira de Ciência da Informação. *In*: PINHO, F. A.; GUIMARÃES, J. A. C. (org.). **Memória, tecnologia e cultura na organização do conhecimento** [online]. Recife, PE: UFPE, 2017. p. 34-47.

GUIMARÃES, J. A. C. Organização do conhecimento: passado, presente e futuro sob a perspectiva da isko. **Informação & Informação**, Londrina, v. 22, n. 2, p. 84-98, 2017.

GUZMÁN-SÁNCHEZ, Maria Victoria. **Patentometría: herramienta para el análisis de oportunidades tecnológicas**. 1999. 130 f. Tese (Doutorado em Gerência de Información Tecnológica) – Facultad de Economía, Universidad de La Habana, Cuba, 1999.

HJØRLAND, B. Citation analysis: a social and dynamic approach to knowledge organization. **Information Processing and Management**, [s. l.], v. 49, n. 6, p. 1313-1325, 2013.

HJØRLAND, B. Domain analysis in information science: eleven approaches traditional as well as innovative. **Journal of Documentation**, [s. l.], v. 58, n. 4, p. 422-462, 2002.

HJØRLAND, B. Domain analysis. **Knowledge Organization**, v. 44, n. 6, p. 436- 464, 2017.

HJØRLAND, B. Domain analysis: A Socio-cognitive orientation for information science research. **Bulletin of the American Society for Information Science and Technology**, [s. l.], v. 30, n. 3, p. 17-21, 2004.

HJØRLAND, B.; ALBRECHTSEN, H. Toward a new horizon in Information Science: domain analysis. **Journal of the American Society for Information Science**, [s. l.], v. 46, n. 6, p. 400-425, jul. 1995.

HUANG, M. H; CHIANG, L. Y.; CHEN, D. Z. Constructing a patent citation map using bibliographic coupling: a study of Taiwan's high-tech companies. **Scientometrics**, Dordrecht, v. 58, n. 3, p. 489-506, 2003.

HUANG, Mu-Hsuan *et al.* Identify technology main paths by adding missing citations using bibliographic coupling and co-citation methods in photovoltaics. *In: PROCEEDINGS OF PICMET '11: TECHNOLOGY MANAGEMENT IN THE ENERGY SMART WORLD (PICMET)*, 2011, Portland. **Anais [...]**.Portland: IEEE, 2011. p. 1-6. Disponível em: <https://ieeexplore.ieee.org/document/6017653>. Acesso em: 31 maio 2022.

INPI. **Estudo comparativo dos critérios de patenteabilidade para invenções biotecnológicas em diferentes países**. Rio de Janeiro: INPI, 2007.

INPI. **Patentes: guia básico**. Rio de Janeiro: INPI, 2020.

IOANNIDIS, J. P. A.; BAAS J.; KLAVANS R.; BOYACK K. W. A standardized citation metrics author database annotated for scientific field. **PLoS Biology**, São Francisco, v. 17, n. 8, aug. 2019. DOI: doi.org/10.1371/journal.pbio.3000384. Disponível em: <https://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.3000384>. Acesso em: 20 maio 2022.

JAPÃO. **Patent Act (Act n° 121 of 1959 as amended in 2006)**. Tokyo: Cabinet Office, Government of Japan, 2010.

JÜRGENS, B.; HERRERO-SOLANA, V. Patent bibliometrics and its use for technology watch. **Journal of Intelligence Studies in Business**, [s. l.], v. 7, n. 2, 2017.

KARKI, M. Patent Citation Analysis: A policy analysis tool. **World Patent Information**, [s. l.], v. 19, n. 4, p. 269-272, 1997.

KESSLER, M. M. An experimental study of bibliographic coupling between technical papers (Corresp.). **IEEE Transactions on Information Theory**, [s. l.], v. 9, n. 1, p. 49-51, 1963.

KESSLER, M. M. Comparison of the results of bibliographic coupling and analytic subject indexing. **American Documentation**, [s. l.], v. 16, n. 3, p. 223-233, 1965.

KUAN, Chung-Huei; HUANG, Mu-Hsuan; CHEN, Dar-Zen. Missing links: Timing characteristics and their implications for capturing contemporaneous technological developments. **Journal of Informetrics**, [s. l.], v. 12, n. 1, p. 259-270, 2018.

KUMARI, R.; JEONG, J.Y.; LEE, B-H.; CHOI, K-N.; CHOI, K. Topic modelling and social network analysis of publications and patents in humanoid robot technology. **Journal of Information Science**, [s. l.], v.47, n. 5, p. 658-676, 2021.

LAI, K.-K.; WU, S.-J. Using the patent co-citation approach to establish a new patent classification system. **Information Processing & Management**, [s. l.], v. 41, n. 2, p. 313-330, 2005.

LEYDESDORFF, L. Patent classifications as indicators of intellectual organization. **Journal of the American Society for Information Science and Technology**, [s. l.], v. 59, n. 10, p. 1582-1597, 2008b.

LEYDESDORFF, L. The delineation of nanoscience and nanotechnology in terms of journals and patents: A most recent update. **Scientometrics**, Dordrecht, v. 76, n. 1, 2008a.

LEYDESDORFF, L.; MEYER, M. The Triple Helix of university-industry government relations. **Scientometrics**, Dordrecht, v. 58, p. 191-203, 2003. Disponível em: https://www.researchgate.net/publication/226095485_Triple_Helix_of_University-Industry-Government_Relations. Acesso em: 22 jan. 2022.

LO, S. C. Patent coupling analysis of primary organizations in genetic engineering research. **Scientometrics**, Dordrecht, v. 74, n.1, p.143–151, 2007. DOI: 10.1007/s11192-008-0110-7. Disponível em: https://ideas.repec.org/a/spr/scient/v74y2008i1d10.1007_s11192-008-0110-7.html. Acesso em: 22 jan. 2022.

LUCAS, E. O.; GARCIA-ZORITA, J. C.; SANZ-CASADO, E. Evolução histórica de investigação em informetria: ponto de vista espanhol. **Liinc em Revista**, Rio de Janeiro, v. 9, n. 1, p. 255-270, 2013.

MACHLUP, F. **An Economic Review of the patente system**: Pursuant to S. Res. 236 Study n.15. Washington: United States Senate, 1958.

MACIAS-CHAPULA, C. O papel da informetria e da cienciometria sua perspectiva nacional e internacional. **Ciência da Informação**, Brasília, DF, v. 27, n. 2, p. 134-140, 1998.

MAGNUS, A. P. M. **Interação entre produção tecnológica e científica**: panorama das patentes e artigos dos pesquisadores dos programas de pós-graduação do instituto de química da UFRGS. 2018. 154 f. Dissertação (Mestrado em Comunicação e Informação) – Universidade Federal do Rio Grande do Sul, Porto Alegre, 2018.

MAI, J-E. Analysis in indexing: document and domain centered approaches. **Information Processing and Management**, [s. l.], v. 41, p. 599-611, 2005.

MARCO, A. C. The dynamics of patent citations. **Economics Letters**, [s. l.], v. 94, p. 290-296, 2007.

MARCONI, M. A.; LAKATOS, E. M. **Fundamentos de metodologia científica**. 6. ed. São Paulo: Atlas, 2009.

MARICATO, J. M. **Dinâmica das relações entre ciência e tecnologia**: estudo Bibliométrico e Cientométrico de múltiplos indicadores de artigos e patentes em biodiesel. 2010. 378 f. Tese (Doutorado em Ciência da Informação) – Programa de Pós-Graduação em Ciência da Informação, Escola de Comunicação e Artes, Universidade de São Paulo, São Paulo, 2010.

MARSHAKOVA, I. V. Citation networks in information science. **Scientometrics**, Dordrecht, v. 3, n. 1, p. 13-25, 1981.

MARTELETO, R. M.; CARVALHO, L. S. Health as a Knowledge Domain and Social Field: Dialogues with Birger Hjørland and Pierre Bourdieu. **Knowledge Organization**, [s. l.], v. 42, p. 581-590, 2015.

MARTÍNEZ, M. J. V. **Análisis de los estudios métricos de la información publicados em revistas españolas de documentación (2005-2009)**. 2011. Licenciatura em Documentación (Proyecto Final de Carrera) – Escuela Técnica Superior de Ingeniería Informática, Universidad Politécnica de Valencia, Valencia, 2011.

MARTÍNEZ-MÉNDEZ, F.; PASTOR-SÁNCHEZ, J. A.; LÓPEZ-CARREÑO, R. Las patentes como indicador de la actividad científica en las universidades españolas. **El profesional de la información**, León, v. 19, n. 2, p. 168-174, mar./abr., 2010.

MATTSON, P. **European Knowledge Transfer Reflected by Research Collaboration and Patent Citations Indicators**. Stockholm: Karolinska Institute, 2011.

MENDEZ-OTERO, R. *et al.* Introdução às células-tronco. In: MORALES, M. M. (ed.). **Terapias avançadas**: células-tronco, terapia gênica e nanotecnologia aplicada à saúde. São Paulo: Atheneu, 2007.

MEYER, M. Tracing knowledge flows in innovation systems. **Scientometrics**, Dordrecht, n. 54, 193-212, 2002.

MEYER, M.; GRANT, K.; MORLACCHI, P. *et al.* Triple Helix indicators as an emergent area of enquiry: a bibliometric perspective. **Scientometrics**, Dordrecht, v. 99, p. 151-174, 2014.

MOED, H. F. **Applied evaluative informetrics**. Dordrecht: Springer, 2017.

MOGEE, M. E.; RICHARD, G. K. Patent citation analysis of Allergan pharmaceutical patents. **Expert Opinion on Therapeutic Patents**, London, v. 8, n. 10, p. 1323-1346, out. 1998.

MOGEE, M. E.; KOLAR, R.G. Patent co-citation analysis of Eli Lilly & Co. **Expert Opinion on Therapeutic Patents**, London, v. 9, n. 3, p. 291–305, mar.1999.

MOURA, A. M. M. **A interação entre artigos e patentes: um estudo cientométrico da comunicação científica e tecnológica em biotecnologia**. 2009. Tese (Doutorado em Comunicação e Informação) – Faculdade de Biblioteconomia e Comunicação – Universidade Federal do Rio Grande do Sul, Porto Alegre, 2009. Disponível em: <https://lume.ufrgs.br/bitstream/handle/10183/18561/000715088.pdf?sequence=1&isAllowed=y>. Acesso em: 10 set. 2022.

NAMBIAR T.S. *et al.* CRISPR-based genome editing through the lens of DNA repair. **Molecular Cell**, New York, v. 82, n. 2, p. 348-388, jan. 2022.

NARIN, F. **Evaluative bibliometrics: the use of publications and citation analysis in the evaluation of scientific activity**. Cherry Hill: Computerhorizons, 1975.

NARIN, F. **Evaluative bibliometrics: the use of publications and citation analysis in the evaluation of scientific activity**. Cherry Hill: Computerhorizons, 1976.

NARIN, F. Patent bibliometrics. **Scientometrics**, Dordrecht, v. 30, p. 147-155, 1994.

NARIN, F.; OLIVASTRO, D.; STEVENS, K. A. Bibliometrics: Theory, practice and problems. **Evaluation Review**, California, v. 18, n. 1, p. 65-76, 1994.

NEIGHBORS, J.M. **Software Construction Using Components**. 1981. 75 p. Tese (Doutorado Filosofia em Ciência da Informação e Computação) - Universidade da Califórnia, Irvine, 1981.

NUNES, G. A. I. O Monopólio como Consequência da Lei de Patentes: as Discordâncias do Direito à Exclusividade. **Revista de Iniciação Científica e Extensão da Faculdade de Direito de Franca**, Franca, v. 5, n. 1, 2020.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT. **Manual de Estatística de Patentes da OECD**. 2. ed. Paris: OECD, 2009.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT. **OECD Patent Statistics Manual**. Paris: OECD, 2009. Disponível em: <https://www.oecd-ilibrary.org/docserver/9789264056442-en.pdf?expires=1654197143&id=id&accname=guest&checksum=05322A137FB83A4BDB0725F242762BD8>. Acesso em: 31 maio 2022.

PEREIRA, C. S. A.; FUJINO, A. A pesquisa sobre patentes na ciência da informação: análise quantitativa da produção científica. *In: ENCONTRO BRASILEIRO DE BIBLIOMETRIA E CIENTOMETRIA*, 4., 2014, Recife. **Anais [...]**. Recife, 2014. p. A47.

PIMENTA, F. P. A patente como fonte de informação (des)necessária para a biotecnologia em saúde. **Transinformação**, Campinas, v. 29, n. 3, 2017.

PRITCHARD, A. Statistical bibliography or bibliometrics. **Journal of Documentation**, v. 25, n. 4, p. 348-349, 1969.

QUONIAM, L. M.; KNISS, C. T.; MAZIERI, M. R. A patente como objeto de pesquisa em ciências da informação e comunicação. **Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação**, Florianópolis, v. 19, n. 39, p. 243-268, 2014. DOI: 10.5007/1518-2924.2014v19n39p243. Disponível em: <https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2014v19n39p243>. Acesso em: 31 maio 2022.

RAVASCHIO, J. P.; FARIA, L. I. L.; QUONIAM, L. M. O uso de patentes como fonte de informação em dissertações e teses de engenharia química: o caso da Unicamp. **Revista Digital de Biblioteconomia & Ciência da Informação**, Campinas, n. 1, v. 8, p. 219-232, 2010.

ROCHA, R. **O direito à vida e a pesquisa em células-tronco**. Rio de Janeiro: Elsevier, 2008.

RODRIGUEZ, A.; KIM, B.; TURKOZ, M. *et al.* New multi-stage similarity measure for calculation of pairwise patent similarity in a patent citation network. **Scientometrics**, Dordrecht, n. 103, p. 565-581, 2015.

ROSTAING, H. **La bibliométrie et ses techniques**. Toulouse: Sciences de la Société, 1996.

SANTOS, F. B. **Produção tecnológica em células-tronco: características e análise de citação das patentes indexadas na base de dados Derwent Innovations Index**. 2018. Dissertação (Mestrado em Comunicação e Informação) – Programa de Pós-graduação em Comunicação e Informação, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2018.

SAYÃO, L. F.; SALES, L. F. **Guia de gestão de dados de pesquisa para bibliotecários e pesquisadores**. Rio de Janeiro: CNEN/IEN, 2015.

SCHULTZ, L. I.; JOUTZ, F. L. Methods for identifying emerging General Purpose Technologies: a case study of nanotechnologies. **Scientometrics**, Dordrecht, v. 85, n. 1, p. 155-170, oct. 2010.

SMIRAGLIA, R. P. Domain analysis of domain analysis for knowledge organization: Observations on an emergent methodological cluster. **Knowledge Organization**, [s. l.], v. 42, n. 8, p. 602-614, 2015.

SMIRAGLIA, R. P. Epistemology of domain analysis. *In: SMIRAGLIA, R. P. LEE, H-L. (ed.). Cultural frames of knowledge*. Germany: Ergon, 2012. p. 111-124.

- SMIRAGLIA, R. P. ISKO 11's Diverse Bookshelf: an editorial. **Knowledge Organization**, [New Delhi], v. 38, n. 3, p. 179-186, 2011.
- SMITH, L. C. Citation Analysis. *In*: POTTER, W.G. **Library Trend**. Champaign: University Illinois, 1981.
- SPINAK, E. **Indicadores cientiométricos de patentes: aplicaciones y limitaciones**. Madrid: [s. n.], 2003.
- STALLMAN, R. **The GNU manifest**. The GNU Project and the Free Software Foundation:1998. Disponível em: <http://www.gnu.org/gnu/manifesto.html>. Acesso em: 02 set. 2022.
- TENNIS, J. T. Two Axes of Domains for Domain Analysis. **Knowledge Organization**, [New Delhi], v. 30, n. 4, 2003.
- TENNIS, J. T. What does a Domain Analysis look like in form, function, and genre? **Brazilian Journal of Information Science**, [s. l.], v. 6, n. 1, 2012.
- TRAJTENBERG, M.; HENDERSON, R.; JAFFE, A. University versus corporate patents: a window on the basicness of invention. **Economics of Innovation and New Technology**, [s. l.], v. 5, n. 1, p. 19-50, 1997.
- VAN ECK, N. J.; WALTMAN, L. **VOSviewer manual**. Leiden. CWTS, 2020. Disponível em : https://www.vosviewer.com/documentation/Manual_VOSviewer_1.6.14.pdf. Acesso em: 14 fev. 2023.
- VAN RAAN, A. F. J. Patent Citations Analysis and Its Value in Research Evaluation: A Review and a New Approach to Map Technology-relevant Research. **Journal of Data and Information Science**, [Beijing], v. 2, n. 1, p. 13-50, 2017.
- VANZ, S. A. A. S.; CAREGNATO, S. N. E. Estudos de citação: uma ferramenta para entender a comunicação científica. **Em Questão**, Porto Alegre, v. 9, n. 2, p. 295-307, 2003.
- VANZ, S. A. S. **A Produção discente em comunicação: análise das citações das dissertações defendidas nos programas de pós-graduação do Rio Grande do Sul**. 2004. 144 f. Dissertação (Mestrado em Comunicação e Informação) – Universidade Federal do Rio Grande do Sul, Porto Alegre, 2004.
- WANG, X.; ZHANG, X.; XU, S. Patent co-citation networks of Fortune 500 companies. **Scientometrics**, Dordrecht, v. 88, p. 761-770, 2011.
- WANG X.; DUAN, Y. Identifying core technology structure of electric vehicle industry through patent Co-citation information. **Energy Procedia**, United Kingdom, v. 5, p. 2581-2585, 2011. DOI: <https://doi.org/10.1016/j.egypro.2011.03.443>.
- WHITE, H. D. Authors as citers over time. **Journal of the American Society for Information Science and Technology**, Nova Jersey, v.52, n.2, p.87-108, 2001.

WHITE, H. D.; GRIFFITH, B. C. Author cocitation: a literature measure of intellectual structure. **Journal of the American Society for Information Science**, [s. l.], v. 32, n. 3, p. 163-171, 1981.

WORLD INTELLECTUAL PROPERTY ORGANIZATION. **International Patent Classification (IPC)**. Geneva, Switzerland: WIPO, [2020]. Disponível em: <https://www.wipo.int/classifications/ipc/ipcpub/>. Acesso em: 2 mar. 2022.

WORLD INTELLECTUAL PROPERTY ORGANIZATION. **WIPO intellectual property handbook**. Switzerland: WIPO, 2004.

WORLD INTELLECTUAL PROPERTY ORGANIZATION. Standard ST.9. *In*: WIPO. **Handbook on Industrial Property Information and Documentation**. Geneva, Switzerland: WIPO, 2013. Disponível em: <chrome-extension://efaidnbnmnibpcjpcglclefindmkaj/https://www.wipo.int/export/sites/www/standards/en/pdf/03-09-01.pdf>. Acesso em: 31 maio 2022

WU, H.C.; CHEN, H. Y.; LEE, K. Y. Unveiling the core technology structure for companies through patent information. **Technological Forecasting and Social Change**, New York, v. 77, p. 1167-1178, 2010. DOI: <https://doi.org/10.1016/j.techfore.2010.03.013>.

ZHAO, D.; STROTMANN, A. Evolution of research activities and intellectual influences in information science 1996–2005: introducing author bibliographic-coupling analysis. **Journal of the American Society for Information Science and Technology**, Nova Jersey, v. 59, n. 13, p. 2070-2086, 2008.

ZHAO, D.; STROTMANN, A. Intellectual structure of stem cell research: a comprehensive author co-citation analysis of a highly collaborative and multidisciplinary field. **Scientometrics**, Dordrecht, v. 87, p. 115-131, 2011.

ZUCOLOTO, G. F.; FREITAS, R. E. **Propriedade Intelectual e aspectos regulatórios em Biotecnologia**. Rio de Janeiro: IPEA, 2013.

**ANEXO A – QUADRO DE MATÉRIAS PATENTEÁVEIS EM BIOTECNOLOGIA
POR PAÍSES**

Matéria Patenteável em Biotecnologia / Países Seleccionados*1		EUA	EUROPA³	JAPÃO	CHINA	ÍNDIA	BRASIL
1	Microorganismos e materiais biológicos associados isolados ou extraídos da natureza: bactéria, fungos - inclusive levedura, algas, protozoários, vírus e líquens	P _{1,3}	P _{1,2,4}	P ₁	P ₁	NP _{1,2}	NP _{1,2,5}
2	Processos que utilizem microorganismos, enzimas ou materiais biológicos associados	P ₂	P _{2,3}	P _{1,3}	P	P _{1,3}	P _{1,3}
3	Microorganismos geneticamente modificados	P ₃	P _{2,4}	P _{1,2}	P ₁	P _{1,2}	P ₁
4	Processo de produção de microrganismos geneticamente modificados (OGM)	P ₂	P _{2,5}	P _{2,3}	P	P ₃	P ₃
5	Animais não humanos e plantas geneticamente modificados (OGM)	P _{4,5}	P _{6,8,17}	P _{4,5}	NP _{2,3}	NP ¹	NP _{5,6}
6	Processo de produção de animais não humanos ou plantas geneticamente modificados (OGM)	P ₂	P ₆	P ₃	P _{2,3}	P _{1,3,6,9}	P _{3,8,9}
7	Processos usados na produção de variedades de planta	P ₂	P ₆	P ₃	P ₃	P _{1,3}	P _{3,9}
8	Variedade Vegetal²	P _{5,6,7}	NP _{7,12}	P ₄	NP ³	NP ¹	NP _{5,6}
9	Processos usados na produção de variedades de animais não humanos	P ₂	P _{6,8}	P ₃	P ₂	P _{1,3,6}	P _{3,8}
10	Variedade Animal (não humano)	P ₄	NP _{6,12}	P ₅	NP ²	NP ¹	NP ⁵
11	Genes ou fragmentos de DNA; RNA e cromossomos de humanos ou outros animais isolados ou extraídos da natureza	P _{9, 12}	P _{9,14,16,17}	P _{10,11}	P ₅	NP _{1,2,4,5}	NP _{1,5}
12	Genes ou fragmentos de DNA; RNA e cromossomos de microorganismo isolados ou extraídos da natureza	P _{9,12}	P ₉	P _{10,11}	P ₅	NP _{1, 2 4,5}	NP _{1,2,5}

13	Genes ou fragmentos de DNA; RNA e cromossomos de planta isolados ou extraídos da natureza	P _{9,12}	P ₉	P _{10,11}	P ₅	NP _{1,2,4,5}	NP _{1,5}
14	Genes ou fragmentos de DNA; RNA e cromossomos de humanos ou outros animais geneticamente modificados	P _{9,12}	P _{9,14,16}	P _{10,11}	P ₅	I _{1,4,5}	NP/P _{5,7,12}
15	Genes ou fragmentos de DNA; RNA e cromossomos de microorganismo geneticamente modificados	P _{9,12}	P ₉	P _{10,11}	P ₅	P _{4,5}	P _{1,2,5,12}
16	Genes ou fragmentos de DNA; RNA e cromossomos de planta geneticamente modificados	P _{9,12}	P ₉	P _{10,11}	P ₅	I _{1,4,5}	P _{5,12}
17	Células-tronco embrionárias	I ₁₀	I _{10, 11,16, 17}	I _{6,8}	NP _{2,6,7}	NP _{1,6}	NP _{5,10}
18	Células–tronco não embrionárias	P ₁₁	I _{11, 16}	P _{6,8}	P _{2,6,7}	NP _{1,6}	NP _{5,10}
19	Célula Humana	P _{11,15}	P _{9, 17}	P _{8,10}	P _{4,6}	NP ¹	NP ₅
20	Célula, Tecido e Órgão Animal (não humano)	P ₄	P _{6,9,17}	P ₅	P ₂	NP ¹	NP ₅
21	Célula, Tecido e Órgão Vegetal	P ₅	P _{6,9}	P ₄	P ₃	NP ¹	NP ⁵
22	Células fusionadas**	P ₁₃	P ₉	P ₁₀	P ₈	I ⁷	P ₁₁
23	Processo de produção de células fusionadas**	P ₂	P _{3,9}	P ₃	P	P ₃	P _{3,11}
24	Seqüências de aminoácidos, polipeptídeos e enzimas isolados ou extraídos da natureza	P _{9,12}	P ₉	P ₁₁	P ₈	NP ₂	NP ¹
25	Seqüências de aminoácidos, polipeptídeos e enzimas geneticamente modificadas	P _{9,12}	P ₉	P ₁₁	P ₈	P _{1,4,5}	P ₁₂
26	Processos de produção de polipeptídeo	P _{2,12}	P ₃	P ₃	P	P ₃	P _{3,4}
27	Anticorpos monoclonais	P ₁₃	P ₉	P ₁₀	P ₈	I ⁷	P ₁₁
28	Método terapêutico; métodos cirúrgicos em seres humanos ou outros animais vivos, de diagnóstico e de tratamento de doenças	P ₁₄	NP	NP	NP	NP ⁸	NP
29	Processo de clonagem de seres humanos ou outros animais	I/P ⁸	NP/I _{16,13}	I/P _{7,8}	NP/P ⁹	NP _{1,9}	NP/P _{7,8}

30	Processo de modificação de identidade genética ou linhagem germinativa de seres humanos ou outros animais	I/P ^{15,16,17}	NP/P ^{8,16,17}	I/P ^{8,9}	NP/P ^{2,4}	NP/I ^{1,9}	NP/P ^{7,8}
31	Uso de embriões humanos ou outros animais	I/P ^{4,18}	I/P ^{15,16}	I/P ^{8,9}	NP/I ^{2,4}	NP/I ^{1,9}	NP/I ^{5,7}
*P = patenteável ; NP = não patenteável; I = não se conseguiu aferir, indefinido ou em discussão							
** por exemplo, hibridoma							
Atualizada em: 11 de outubro de 2013							

APÊNDICE A – CÓDIGO DOS PAÍSES ONDE AS PATENTES SÃO DEPOSITADAS

Código	Name
AP	Organização Regional Africana da Propriedade Industrial (ARIPO)
AR	Argentina
AT	Áustria
AU	Austrália
BA	Bósnia e Herzegovina
BE	Bélgica
BG	Bulgária
BR	Brazil
CA	Canadá
CH	Suíça
CN	China
CS	Checoslováquia (até 1993)
CU	Cuba
CY	Chipre
CZ	República Checa
DD	Alemanha, excluindo o território que, antes de 3 de Outubro de 1990, constituía a República Federal da Alemanha
DE	Alemanha
DK	Dinamarca
DZ	Argélia
EA	Organização Euroasiática de Patentes
EE	Estónia
EG	Egipto
EP	Organização Europeia de Patentes (OPE/EPO)
ES	Espanha
FI	Finlândia
FR	França
GB	Reino Unido
GR	Grécia
HK	Hong Kong
HR	Croácia
HU	Hungria
IE	Irlanda
IL	Israel
IN	Índia
IT	Itália
JP	Japão
KE	Quênia
KR	República da Coreia
LT	Lituânia
LU	Luxemburgo

LV	Letónia
MC	Mónaco
MD	República da Moldávia
MN	Mongólia
MT	Malta
MW	Malawi
MX	México
MY	Malásia
NC	Nova Caledónia
NL	Holanda
NO	Noruega
NZ	Nova Zelândia
OA	Organização Africana da Propriedade Intelectual (OAPI)
PH	Filipinas
PL	Polónia
PT	Portugal
RO	Roméia
RU	Federação Russa
SE	Suécia
SG	Singapura
SI	Eslovénia
SK	Eslováquia
SU	União das Repúblicas Socialistas Soviéticas (URSS)
TJ	Tadjiquistão
TR	Turquia
TT	Trindade e Tobago
TW	Taiwan
US	Estados Unidos da América
VN	Vietname
WO	Organização Mundial da Propriedade Intelectual (OMPI/WIPO)
YU	Jugoslávia
ZA	África do Sul
ZM	Zâmbia
ZW	Zimbabwe

**APÊNDICE B – TABELA CONTENDO AS 144 PATENTES ANALISADAS COM
SUAS RESPECTIVAS CITAÇÕES E NÚMERO DE ARTIGOS E PATENTES
CITADAS.**

Número	Cit. Recebidas	Art. Cit./ examinador	Art. Cit./ inventor	Total de Art. Citados	Pat. Cit./ examinador	Pat. Cit./ inventor	Total de Patentes citadas
EP3009511	297	36	1104	1140	15	666	681
CN105748343	152	2	0	2	6	0	6
CN105384825	91	15	44	59	26	58	84
WO2016028896	86	30	690	720	4	968	972
US20160143961	70	70	742	812	16	112	128
WO2016049230	65	25	224	249	10	8	18
WO2016205613	64	10	532	542	12	183	195
WO2016126608	62	8	362	370	10	399	409
WO2016106236	62	4	220	224	1	195	196
US20160145646	59	48	292	340	17	333	350
WO2016014576	59	34	700	734	18	952	970
WO2017191274	57	5	13	18	6	106	112
WO2016090320	55	4	296	300	19	88	107
WO2016141224	54	25	421	446	11	341	352
WO2017079673	53	16	95	111	18	121	139
WO2016183041	51	6	16	22	12	50	62
WO2015191693	51	14	30	44	13	26	39
WO2017030814	50	6	112	118	5	142	147
WO2016164356	49	6	147	153	18	52	70
WO2017186718	45	18	33	51	6	0	6
WO2016164731	45	19	694	713	19	738	757
WO2016094880	44	5	249	254	6	141	147
WO2017049166	43	12	0	12	3	0	3
WO2016170176	43	78	293	371	41	195	236
WO2016057961	42	6	47	53	4	2	6
WO2016072821	41	14	0	14	36	13	49
WO2017161371	40	4	5	9	16	6	22
WO2016182959	40	13	93	106	9	5	14
WO2016115326	39	14	0	14	2	0	2
WO2017149515	38	2	183	185	15	290	305
WO2016057821	38	0	0	0	8	0	8
EP3034092	37	6	4	10	1	0	1
CN104920227	37	1	0	1	1	0	1
WO2016019144	36	3	114	117	7	193	200
CN105647968	35	1	0	1	4	0	4
WO2019051380	34	3	70	73	3	0	3
WO2017223330	34	2	2	4	8	2	10
WO2017075537	34	19	209	228	20	95	115
WO2016210447	34	6	64	70	10	10	20
WO2016092106	34	3	2	5	15	4	19
WO2016205759	33	3	0	3	3	0	3
WO2016166340	33	12	209	221	2	29	31
WO2016161516	33	24	276	300	6	5	11
WO2016004875	33	8	0	8	11	0	11
US20170016027	32	25	57	82	14	112	126
WO2017189308	31	0	0	0	0	0	0
WO2017106657	31	0	0	0	0	0	0
WO2017078807	31	17	188	205	21	224	245
WO2016191756	31	13	33	46	8	24	32
WO2018102795	30	4	0	4	7	0	7
WO20161545	30	1	130	131	13	4	17

WO2016138034	30	46	317	363	22	192	214
WO2016086161	30	7	0	7	10	0	10
EP2990416	30	14	51	65	11	48	59
WO2016164937	29	6	306	312	21	267	288
WO2016154055	29	1	0	1	13	0	13
WO2016089866	29	4	0	4	8	0	8
WO2016057835	29	16	0	16	2	0	2
WO2017023801	28	7	352	359	14	546	560
WO2016176191	28	10	53	63	3	36	39
US20160108369	28	37	204	241	8	221	229
WO2016036754	28	10	5	15	9	7	16
EP2985285	28	5	1	6	5	2	7
US20160024474	28	5	0	5	6	0	6
WO2017106528	27	3	0	3	9	0	9
WO2017025323	27	11	88	99	12	51	63
WO2016154579	27	8	32	40	5	1	6
WO2017115268	26	4	35	39	11	14	25
WO2016123578	26	33	88	121	4	33	37
CN105477016	26	7	0	7	1	0	1
WO2016025469	26	16	0	16	10	0	10
WO2016025759	26	5	88	93	27	79	106
WO2018132783	25	3	39	42	5	4	9
WO2018035388	25	7	0	7	17	0	17
WO2017177137	25	8	0	8	31	0	31
WO2017053312	25	3	54	57	7	20	27
WO2017048193	25	12	0	12	3	0	3
WO2016135558	25	22	94	116	8	20	28
WO2016069788	25	4	86	90	23	5	28
WO2016044416	25	45	67	112	17	104	121
WO2017143042	24	21	490	511	12	176	188
WO2017136794	24	3	0	3	4	0	4
WO2017053729	24	4	3	7	7	40	47
CN106367386	24	0	0	0	3	0	3
WO2016201047	24	6	83	89	5	8	13
WO2016191684	24	6	0	6	6	0	6
WO2016141137	24	11	20	31	18	17	35
GB2531454	24	15	295	310	4	307	311
WO2016049251	24	18	203	221	1	192	193
WO2016021972	24	12	17	29	8	14	22
WO2019025984	23	5	0	5	7	0	7
WO2018005873	23	4	264	268	0	116	116
WO2017165464	23	18	91	109	10	142	152
WO2017075389	23	13	85	98	12	21	33
WO2017059168	23	50	330	380	63	679	742
WO2016179242	23	7	58	65	13	20	33
WO2016109668	23	10	0	10	3	0	3
WO2016073955	23	11	0	11	5	0	5
CN105543313	23	1	0	1	12	0	12
WO2016057671	23	13	10	23	1	1	2
CN105331586	23	6	0	6	6	0	6
WO2016016894	23	17	214	231	5	75	80
WO2018121712	22	4	0	4	10	0	10
WO2018073391	22	3	76	79	3	27	30
WO2018064640	22	2	0	2	8	0	8
WO2017011721	22	7	91	98	7	32	39
WO2016142719	22	6	26	32	0	0	0
WO2016026854	22	26	46	72	16	5	21
WO2018208728	21	1	0	1	8	0	8
WO2018195545	21	1	0	1	3	0	3

WO2018112363	21	2	35	37	3	7	10
WO2018035423	21	2	0	2	10	0	10
CN106282107	21	0	0	0	0	0	0
CN106267341	21	0	0	0	5	0	5
WO2016054155	21	0	0	0	8	0	8
WO2016004043	21	8	0	8	18	0	18
WO2018195555	20	4	72	76	3	19	22
CN107723275	20	9	0	9	7	0	7
WO2017109757	20	8	162	170	7	146	153
US9636525	20	0	49	49	0	37	37
WO2017070169	20	5	0	5	5	0	5
CN106474155	20	2	0	2	3	0	3
WO2017024255	20	27	89	116	10	103	113
WO2017004279	20	1	0	1	8	0	8
WO2016142674	20	26	174	200	12	310	322
WO2016135557	20	11	132	143	4	41	45
WO2016135559	20	9	123	132	4	135	139
WO2016100856	20	11	64	75	7	34	41
WO2016040858	20	3	140	143	10	135	145
WO2016025454	20	21	0	21	9	0	9
CN105267240	20	4	0	4	4	0	4
CN105154473	20	12	0	12	8	0	8
WO2017161010	19	7	0	7	10	0	10
WO2017099494	19	4	0	4	1	0	1
WO2017096248	19	7	0	7	8	0	8
WO2017072590	19	4	129	133	3	147	150
WO2017062754	19	3	4	7	4	2	6
WO2016191545	19	11	0	11	8	0	8
WO2016187226	19	16	40	56	7	6	13
WO2016181357	19	3	4	7	4	4	8
WO2016183236	19	5	37	42	13	6	19
CN105769740	19	0	0	0	3	0	3
WO2016081924	19	2	0	2	4	0	4
WO2016075662	19	10	24	34	11	11	22

*Art. Cit./ ex. = Artigo citado pelo examinador; Art. Cit./ inv.= Artigo citado pelo inventor; Pat. Cit./ ex. = patentes citadas pelo examinador; Pat. Cit./ inv. = patentes citadas pelo inventor

Fonte: elaborado pelo autor

APÊNDICE C – LINKS DE ACOPLAMENTO ENTRE PATENTES

Tabela identificando os links que representam a quantidade de patentes que cada patente está acoplada.

id	label	weight	id	label	weight	id	label	weight
1	CN104920227	0	51	WO2016073955	9	101	WO2017030814	1
2	CN105154473	0	52	WO2016075662	3	102	WO2017048193	0
3	CN105267240	0	53	WO2016081924	4	103	WO2017049166	4
4	CN105331586	1	54	WO2016086161	1	104	WO2017053312	24
5	CN105384825	15	55	WO2016089866	11	105	WO2017053729	11
6	CN105477016	0	56	WO2016090320	21	106	WO2017059168	17
7	CN105543313	1	57	WO2016092106	0	107	WO2017062754	6
8	CN105647968	0	58	WO2016094880	47	108	WO2017070169	1
9	CN105748343	0	59	WO2016100856	0	109	WO2017072590	24
10	CN105769740	1	60	WO2016106236	45	110	WO2017075389	5
11	CN106267341	0	61	WO2016109668	8	111	WO2017075537	17
12	CN106282107	0	62	WO2016115326	2	112	WO2017078807	34
13	CN106367386	0	63	WO2016123578	15	113	WO2017079673	33
14	CN106474155	0	64	WO2016126608	33	114	WO2017096248	0
15	CN107723275	7	65	WO2016135557	34	115	WO2017099494	0
16	EP2985285	1	66	WO2016135558	32	116	WO2017106528	2
17	EP2990416	16	67	WO2016135559	26	117	WO2017106657	0
18	EP3009511	55	68	WO2016138034	14	118	WO2017109757	24
19	EP3034092	0	69	WO2016141137	1	119	WO2017115268	4
20	GB2531454	39	70	WO2016141224	36	120	WO2017136794	11
21	US2016024474	0	71	WO2016142674	0	121	WO2017143042	30
22	US2016108369	41	72	WO2016142719	0	122	WO2017149515	29
23	US2016143961	1	73	WO2016154055	1	123	WO2017161010	4
24	US2016145646	61	74	WO2016154579	21	124	WO2017161371	0
25	US2017016027	18	75	WO2016154596	19	125	WO2017165464	7
26	US9636525	0	76	WO2016161516	1	126	WO2017177137	12
27	WO2015191693	31	77	WO2016164356	32	127	WO2017186718	5
28	WO2016004043	0	78	WO2016164731	33	128	WO2017189308	0
29	WO2016004875	0	79	WO2016164937	6	129	WO2017191274	1
30	WO2016014576	49	80	WO2016166340	18	130	WO2017223330	0
31	WO2016016894	2	81	WO2016170176	7	131	WO2018005873	36
32	WO2016019144	29	82	WO2016176191	12	132	WO2018035388	22
33	WO2016021972	14	83	WO2016179242	1	133	WO2018035423	15
34	WO2016025454	7	84	WO2016181357	2	134	WO2018064640	0
35	WO2016025469	7	85	WO2016182959	19	135	WO2018073391	16
36	WO2016025759	37	86	WO2016183041	25	136	WO2018102795	7
37	WO2016026854	0	87	WO2016183236	19	137	WO2018112363	1
38	WO2016028896	48	88	WO2016187226	1	138	WO2018121712	9
39	WO2016036754	21	89	WO2016191545	0	139	WO2018132783	5
40	WO2016040858	0	90	WO2016191684	15	140	WO2018195545	3
41	WO2016044416	27	91	WO2016191756	9	141	WO2018195555	9
42	WO2016049230	1	92	WO2016201047	29	142	WO2018208728	0
43	WO2016049251	38	93	WO2016205613	46	143	WO2019025984	0
44	WO2016054155	0	94	WO2016205759	10	144	WO2019051380	0
45	WO2016057671	0	95	WO2016210447	13			
46	WO2016057821	17	96	WO2017004279	30			
47	WO2016057835	13	97	WO2017011721	20			
48	WO2016057961	15	98	WO2017023801	73			
49	WO2016069788	0	99	WO2017024255	0			
50	WO2016072821	0	100	WO2017025323	22			

Est. Descritiva	
Mínimo	0
Máximo	73
Média	12,97
Desv. Padrão	15,31
Coef. Variação	1,18
Ligações Totais (Acopl. Totais)	934
Patentes	144
Ligações Possíveis	10296
Densidade	9,07%

Fonte: elaborado pela autora a partir da rede criada no VOSviewer.