Patent search strategies in plant biotechnology

Claudia Parisi (✉ claudia.parisi@ec.europa.eu)

European Commission Joint Research Centre - IPTS

Emilio Rodríguez-Cerezo

European Commission Joint Research Centre - IPTS

Harry Thangaraj

St George's University London

Method Article

Abstract

This protocol provides basic guidelines for performing a patent search through illustration with a specific biotechnology subject. It is directed in particular to beginners who wish to broaden their knowledge by using patents as a source of information. The instructions given show how to search for patents claiming the use of ZFN (Zinc Finger Nucleases) in plants, a recently developed gene targeting technique with wide potential in plant breeding. The guidelines focus in particular on the use of two online available patent databases: esp@cenet, from the European Patent Office, and PatentScope, from the World Intellectual Property Organization. Several exercises are shown as examples to follow in searches, in particular the use of keywords and patent codes. We believe that the strategies proposed in this protocol are very useful for any type of search, but each searcher can choose a method she or he feels more comfortable with.

Introduction

This protocol provides basic step-by-step instructions to build a patent search methodology on a specific scientific subject. The protocol refers to the search for patents related to technologies that involve targeted mutagenesis and genetic manipulation in plants using zinc finger nucleases [1] as a tool. Due to the innovative potential of this plant breeding technique, it is of interest to explore the patent landscape related to this subject, as was done by Lusser et al [2]. The protocol illustrates the use of the two online patent databases esp@cenet, administered by the European Patent Office (EPO), and PatentScope, administered by the Word Intellectual Property Organisation (WIPO). We consider them as the most valuable freely available databases, in terms of content and functionality. esp@cenet includes not only European patents, but also patents administered by WIPO and lastly the "worldwide" subset: patents filed in over 80 countries. Its search functions allow to search for specific words in different fields of the patents, like title, abstract, full text (only for European or WIPO patents), inventor field, assignee field etc. PatentScope contains all patents that were submitted through the PCT route [3], and possess similar search functionalities as esp@cenet, and in addition the possibility of search in the field of patent claims. This function is very useful, since claims define precisely the scope of protection of each patent. Searching through keywords in specific fields, and also through patent codes is recommended and each described in this protocol. Patent examiners assign to each patent one or more codes that classify technology areas according to a hierarchical system. They follow the International Patent Classification (IPC) system. esp@cenet also employs the European classification system (ECLA), which is more detailed, drilling to lower levels of hierarchy and making the search more precise. The protocol represents a proposed method to use all the described tools in the search of patents related to ZFN in plants, but we want to highlight that several other strategies can be employed, depending on the subject to be searched or on the preferences of the searcher.

Equipment

The search only requires a computer with access to Internet, since the patent databases we suggest for the search are publicly available online for free.

Procedure

1) BACKGROUND RESEARCH: PRELIMINARY SEARCHES IN SCIENTIFIC LITERATURE AND INDUSTRY JOURNALS **1.1 Literature search** Perform a literature search on ZFN techniques in databases like ISI Web of Knowledge, PubMed or ScienceDirect in order to retrieve information about the technique and its application and in order to discover useful words to be used in the patent search. Use the query "zinc finger nuclease**" or ZFN for a general search. Add to the previous queries "AND plant**" or "AND crop*" or the name of specific plants like maize, tobacco, Arabidopsis, etc. for a search directed to publications about plants. From the publications retrieved, note also the names of authors and affiliations, since they can correspond to patent inventors and assignees. **1.2. Search in industry journals** Search also for more information in industry website like the website of the magazine Genetic Engineering & Biotechnology News \("http://www.genengnews.com/":http://www.genengnews.com/) by using the same words in the "search" fields. You will retrieve more information about the interest by the industry on this technique. **2) SEARCHING WITH KEYWORDS IN ESP@CENET** **2.1. Starting with a simple search** Go to "http://worldwide.espacenet.com/advancedSearch?locale=en_EP":http://worldwide.espacenet.com/advancedSearch?locale=en_EP. In the section "1. Database", you can choose the specific subset to which to direct your search \(European patents, PCT patents or worldwide), while in the section "2. Search terms" you can introduce the keywords for the search in specific fields. Perform an initial search for patents that might precisely correspond to the subject of interest, for example introducing the most meaningful search words in the field title or abstract, like in the example: • Keyword\(s) in title or abstract: \("zinc finger nuclease?" OR ZFN) AND plant? 6 results in the worldwide subset Many of the findings are very useful examples of inventions on ZFN technique in plants, although not comprehensive. You can analyse them, learning about their content and improve your strategy, e.g. by indentifying further keywords. **2.2. Experimenting with more search words** Try to use the queries of the previous exercise, within the whole text of the patent description \(in WIPO or EPO subsets): • Keyword\(s) in full text: \("zinc finger nuclease?" OR ZFN) AND plant? 129 results in the WIPO subset 8 results in the EPO subset You will clearly obtain more results than previously. This is because the patent description is generally a very complex text and both "zinc finger nuclease" and "plants" might appear in the full description of an invention even if the main concept is not related to the use of ZFNs in plants. Thus for a search within the whole text it may be advisable to expand search queries with additional keywords and Boolean logic, in order to restrict the number of results to a more "browsable" list, like in the examples: • Keyword\(s) in full text: "zinc finger" AND nuclease? AND plant? AND "double strand break" 58 results in the WIPO subset 5 results in the EPO subset • Keyword\(s) in full text: \("zinc finger protein" OR "zinc finger nuclease") AND plant? AND "double strand break" 37 results in the WIPO subset 4 results in the EPO subset • Keyword\(s) in full text: "zinc finger" and nuclease and "non homologous end joining" and plant? 9 results in the WIPO subset 2 result in the EPO subset Consider the following search words or phrases for use in further queries: zinc finger protein, ZFP, chimeric nuclease,

double strand break, zinc finger binding domain, DNA double strand break inducing activity, cleavage domain \(or half-domain), targeted cleavage, homologous recombination, non-homologous end joining, site- specific modification, targeting vector, recognition sequence, insertion sequence, transgene integration, plant host cell, etc. Screen manually the list of results to identify patents that reflect the objective of your search. esp@cenet offer different options for storing and screening results, like e.g. the creation of a patent list or the export of the results to an excel sheet. **2.3 Restricting the results** You can consider the use of the Boolean operator NOT to exclude a group of patents from the list. For example, ZFNs are frequently applied in research to human cells for the correction of genetic disorders caused by point mutations. If you want to only focus to plant breeding and wish to clean up your findings from these patents, try using NOT accompanied by words like "human disease" or "human disorders" or cancer etc., as illustrated below: • Keyword\(s) in full text: "zinc finger" AND nuclease? AND plant? AND "double strand break" NOT \("human disease?" OR cancer) 32 results in the WIPO subset 3 result in the EPO subset Caution is required\! Excluding patents containing certain phrases or other types of queries may also result in over-filtration and loss of relevant patent information. **2.4 Search using inventors' and assignees' names** Introduce general keywords in full text and limit the search to potential assignees \(field "Applicant\(s)") or inventors \(field "Inventor\(s)") that you have identified in your previous searches \(including information derived from literature and industry journals). Examples: • Keyword\(s) in full text: "zinc finger nuclease?" AND plant? Applicant\(s): sangamo OR dow OR sigma OR "univ iowa" OR "univ michigan" OR "univ utah" 40 results in the WIPO subset 4 results in the EPO subset • Keyword\(s) in title or abstract: "zinc finger nuclease?" Applicant\(s): sangamo OR dow OR sigma OR "univ iowa" OR "univ michigan" OR "univ utah" 76 results in the worldwide subset • Keyword\(s) in full text: "zinc finger" Inventor\(s): urnov OR petolino OR shukla OR carroll OR miller 50 results in the WIPO subset 8 results in the EPO subset **3) USING PATENT CODES IN ESP@CENET** In esp@cenet, choose the page of "classification search": http://worldwide.espacenet.com/eclasrch?locale=en_EP&classification=ecla Click sequentially on each code: the description of the category and all relative subgroups will appear. Search for codes that might be assigned to some patents related to zinc finger nucleases. Examples: C12N15/90 \(or on "stable integration of foreign gene into chromosome") C12N15/82A10 \(\[N: Targeted insertion of genes into the plant genome by homologous recombination]) To "extract" more codes that might pertain to the underlying technology, observe the list of findings obtained in the previous exercises. You will note from the results list that some of the most recent entries have only IPC codes assigned to them. The older patents have ECLA codes. Choose "ECLA" to sort the results by patent codes. Looking at the results set carefully will give a good idea of the patent codes assigned by examiners to this area of technology, which will help you expand your searches considerably. Look through the list and you will come across for example with ECLA code "C12N15/82A10". **3.1 Only using codes** You will notice that many of the codes relate to broad areas of technology. Obviously, using these broad codes in isolation will give you very broad result lists. Examples: • Keyword\(s) in European Classification \(ECLA): C12N15/82A10 1,119 results in the worldwide subset • Keyword\(s) in European Classification \(ECLA): C12N15/90B 864 results in the worldwide subset **3.2 Combining codes with words queries** In order to narrow down the list to more specific results you have to judiciously use patent codes in combination with word queries, and also experiment with different fields, like in the following examples: • Keyword\(s)

in title or abstract: "zinc" Keyword\(s) in European Classification \(ECLA): C12N15/82A10 9 results in the worldwide subset • Keyword\(s) in title or abstract: "zinc" Keyword\(s) in European Classification \(ECLA): C12N15/90B 3 results in the worldwide subset • Keyword\(s) in full text: "zinc finger nuclease?" Keyword\(s) in International Patent Classification \(IPC): C12N15 9 results in the EPO subset 93 results in the WIPO subset • Keyword\(s) in full text: "zinc finger nuclease?" and plant? Keyword\(s) in International Patent Classification \(IPC): C12N15 8 results in the EPO subset 84 results in the WIPO subset Full text searching is not available for the worldwide collection subset. Therefore, you can combine a word with the broader IPC code C12N15. As with all exercises and results, you should manually sift through the results to assess how relevant each result is, as illustrated in the previous chapters. **3.3 Further exercises** Try using the search queries indicated in the previous exercises, combined with the identified codes. Results might not be satisfactory for ZFN technology, but the scope of this exercise is to give instructions that can be applied to many other inventive fields in biotechnology. Given the variety of IPC and ECLA codes available, other inventive fields could be more precisely represented by their relative codes, than the example of ZFN technology, as yet there is no precise code that relates to ZFN in the IPC or ECLA. In contrast some commercial databases have their own manually assigned codes for ZFN \(for example – Derwent World Patent Index, DWPI, has code B04-N11 for zinc finger proteins). **4) PATENT SEARCH IN PATENTSCOPE DATABASE**: The next set of examples is designed to gain initial familiarity with the PatentScope database \(
"http://www.wipo.int/pctdb/en/index.jsp":http://www.wipo.int/pctdb/en/index.jsp) maintained by WIPO. **4.1 Simple search in PatentScope** To access the search page for a simple search go to: http://www.wipo.int/pctdb/en/search-simp.jsp. Make sure that only "PCT" is ticked, and uncheck any other boxes. Add a search query in the text box and make sure that the "front page" tab is selected \(taking note that a number of other tabs allow us to search other fields, such as IPC classification), like in the examples: • Keyword\(s) in "front page" - PCT: "zinc finger nuclease?" 10 results in PatentScope • Keyword\(s) in "front page" - PCT: zinc AND finger AND nuclease? 13 results in PatentScope • Keyword\(s) in "any field" - PCT: "zinc finger nuclease?" 130 results in PatentScope Additionally you can try searching without the quotes. The front page field usually includes the title and abstract entries. By choosing the "any field" option in the banner, the search is performed in the entire text of the patent. **4.2 Combining fields**: A more useful approach is to choose the "search" option in the header of the page and choose the "field combination" option. This takes you to a page that helps you combine various fields in the searches: http://www.wipo.int/patentscope/search/en/structuredSearch.jsf On the left, through drop down lists, it is possible to change the default fields shown. This makes it possible to combine same or different fields. Also it is possible to combine the fields through either Boolean operators AND /OR. Again make sure that only the PCT option is selected, rather than national collections. Amongst the fields, "International Class" will allow you to enter IPC codes, although it may be possible that lower level hierarchy codes or the more detailed codes will not work. Here you can experiment, as in the example: • Keyword\(s) in "any field": "zinc finger nuclease?" 130 results in PatentScope Then you can redo the same search but by adding "plant" to the "English claims" field, which narrows down the results: • Keyword\(s) in "any field": "zinc finger nuclease?" Keyword\(s) in "English claims": plant? 7 results in PatentScope You can also combine search queries and codes, as you did for

esp@cenet. Select the field "International Class" and insert the codes identified previously, like C12N-15/90 \(note this format is somewhat different from ECLA code shown in an earlier example when searching esp@cenet). You can combine that with "plant" in the claim field: • Keyword\(s) in "International Class": C12N-15/90 Keyword\(s) in "English claims": plant? 29 results in PatentScope • Keyword\(s) in "International Class": C12N-15/82 Keyword\(s) in "English claims": plant? 2,526 results in PatentScope The records will most likely cover some means of integrating DNA into the chromosome – and combining it with "plant" in claims would mean all of these include the word "plant" in at least one of the claims. **4.3 Combining claims** You can select twice, in two different drop down boxes, the field "English claim" and introduce different queries like in the example: • Keyword\(s) in " English claims ": "zinc finger" Keyword\(s) in "English claims": plant? 35 results in PatentScope In this case, a similar result would be obtained by using the Boolean operator AND in the same field. Searching the word "plant" in the claims would likely pick up patents where the claimed technology is also directed to the use in plants. But you have to consider the possibility that the claims may include a word such as "Arabidopsis" or "Zea Mays" instead of "plant." Therefore, one can also try making the searches previously illustrated, but substituting the word "plant" with specific plant names in the queries.

Anticipated Results

Fig. 1 shows the results obtained about patents \(both applications and granted patents) on ZFN in plants, following the described protocol.

References

1 Shukla, V. K. et al. Precise genome modification in the crop species Zea mays using zinc-finger nucleases. Nature 459, 437-441 \(2009). 2 Lusser, M., Parisi, C., Plan, D. & Rodriguez-Cerezo, E. Deployment of new biotechnologies in plant breeding. Nat Biotech 30, 231-239 \(2012). 3 WIPO. Patent Cooperation Treaty. Done at Washington on June 19, 1970, amended on September 28, 1979, modified on February 3, 1984, and on October 3, 2001. http://www.wipo.int/pct/en/texts/articles/atoc.htm. \(2001).
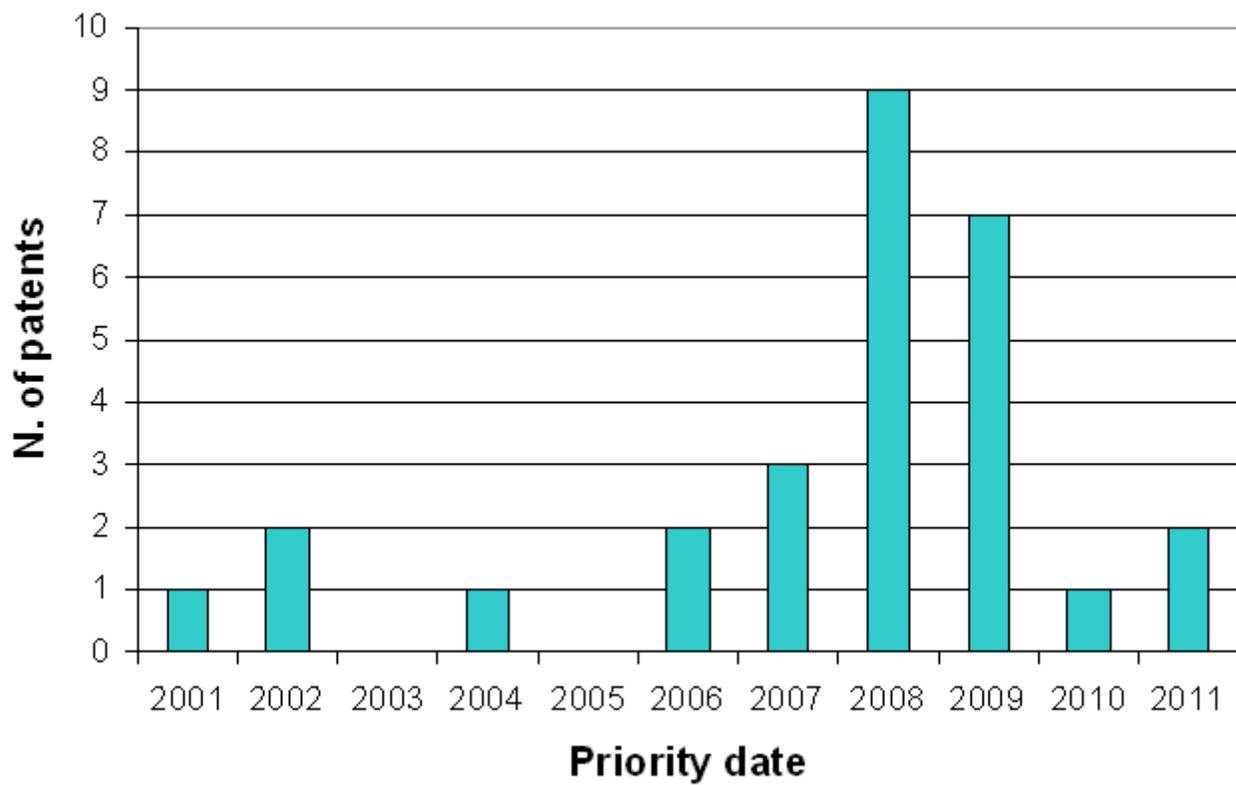
Acknowledgements

Figures

Figure 1

Patent results on the use of ZFN in plants, distributed according to priority date Both granted patents and patent applications have been included in the count and each patent represents all the members of its patent family.