

Patents on genetic sequences: excess and fragility

Description

In its quest for the ultimate protection, the “life sciences” industry has invented ways of drafting its patents that are more or less legitimate and solid. Here we look at the main ones.

Genetic sequences (of DNA or RNA [1]) are a succession of “*letters*” (or nucleic basis or nucleotides) whose translation produces a protein. The first patent on a genetic sequence was granted in 1982 for a human growth hormone. Already, the drafting of claims [2] of this patent shows the complexity of the exercise [3]. Starting with a truly identified and unique sequence, applicants try, *via* the play of wording, to cover all possible variants. For fear of missing out on optimum protection, their claims are far too ambitious in scope. This is often unjustified, because not all the sequences covered by the claims are useful or functional.

Mastering the “*timing*” and writing style

Patent applications covering genetic sequences or proteins are filed as early as possible, with claims drafted as broadly as possible. But often, the industrial application of the sequences, or their “*utility*” if we take the notion of US law, is hypothetical. However, the EPO (European Patent Office) requires that “*where it is not obvious from the description or the nature of the invention, the manner in which [the invention] is susceptible of industrial application*” (Rule 42 (1) f EPC) be explained [4]. This rule clearly states that a simple DNA (or RNA, or protein) sequence without any indication of function cannot be patented.

At the application stage, the drafting objective for the claims is limited to fulfilling the “*novelty*” requirement [5]. In other words, define an invention that is not already described in “prior art”. If broad protection of the invention is sought, the invention must be sufficiently supported in all its scope by various embodiments. A claim must also be “*clear*”, “*concise*” and “*based on the description [of the invention]*” (you cannot claim more than the detailed description of the invention in the patent application) [6].

The “*biotech*” sector has therefore sought to develop specific drafting techniques to accommodate all of these legal requirements. An exhaustive study of all these techniques is not proposed here. The most common, among those that are intended to be the most ‘avid’, are nevertheless discussed. These techniques all have one thing in common: they seek to protect genetic sequences on a very large scale beyond those that were actually identified at the outset.

Pretends of identity

One approach is to use the concept of “*percent identity*”. In practical terms, this means that any sequence which is, for example, at least 80%, 90% or 95% identical to the sequence actually “*invented*” is also covered by the patent. The lower the percentage, the greater the range of sequences covered by the patent. Even with 95%, this still corresponds to a great diversity of sequences, which are of

course not all individually described by the patent. Often, only a few examples are provided to create the illusion that the scope of the protection claimed has been explored and is therefore justified. However, it is very plausible that many of the sequences (or “*variants*”) claimed are not in fact functional and therefore cannot meet the conditions for patentability. A famous British judge and patent lawyer, N.R. Pumfrey, spoke of a “*armchair invention*” that could not be implemented to its full extent. To exclude such non-functional sequences, the claims specify that these “*variants*” must have the same biological activity as the sequence actually identified.

In a minority of cases, patent applications replace this notion of “*identity*” with that of “*homology*” or “*similarity*”. In simple bioinformatics language, two genetic sequences are said to be “*identical*” if they have exactly the same alignment of nucleotides (the “*letters*” of DNA). They are “*similar*” if they are relatively close but not identical. Homology is a special case of similarity in which two similar sequences are considered to derive from a common ancestor, i.e. the sequences are evolutions of a common initial sequence in which certain parts have been added, deleted or replaced [7]. The European Patent Office provides its interpretation of these terms in its Guidelines [8].

In order to get an idea of the volume of patents granted using this type of language, we conducted a search *via* the Lens tool on European and US patents that deal with sequences using the keywords “*identity*”, “*homology*” and “*similarity*” [9]. One explanation for this is that the vast majority of applicants are American and have not extended their patents to Europe. Another could be that the EPO is less inclined than the USPTO (US Patent & Trademark Office) to issue this type of claim. Nevertheless, it does so (see box) and, to this end, proposes guidelines to applicants [10].

In Brazil, the Biotechnological Inventions Guidelines, amended on 1^{er} December 2020, clearly address the issue: “*A claim with terms of identity or similarity cannot be accepted as the manner in which it is worded encompasses countless different sequences, not even specifying at what locations on the nucleotides (or amino acid) sequence such substitutions might occur; consequently, claims of this type may not be accepted, as the characterization of the object of the protection is not clear and precise, contravening article 25 of the Brazilian IP Statute [11]*”.

[Example of European claim using the concept of sequence identity](#)

European patent EP3720955B1: “*Variants of alpha-amylase and polynucleotides encoding them*” granted on 14 June 2023 to Novozymes (DK) – claim 1 (partial wording)

1. An alpha-amylase variant comprising a substitution at a position corresponding to position 188 and at least one additional substitution at a position corresponding to position 242 or 279 or 275 of SEQ ID [12] NO: 1, the substitutions being selected from the group consisting of E188P+S242Y, [...] and E188P+N275H, the variant having at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99%, but less than 100% sequence identity with a parent alpha amylase selected from the group consisting of [...].

“Functional” language or the cult of vagueness

An applicant can, if his “*invention*” lends itself to it, aim even wider. By avoiding defining a sequence in the wording of a claim, he does not limit the invention to any particular sequence. An invention can in fact be defined in purely functional terms, including in the case of genetic sequences. This is what the EPO says in language that may be a little convoluted: “A claim may broadly define a feature in terms of its function, i.e. as a functional feature, even where only one example of the feature has been given in the description, if the skilled person would appreciate that other means could be used for the same function” [13].

However, the Guidelines state: “[...] it may not be sufficient if the description merely states in vague terms that other means may be adopted, if it is not reasonably clear what they might be or how they might be used.”. In other words, the description of the patent must be sufficient to enable “the person skilled in the art” to determine the various means that can be used to implement an invention defined in purely functional terms.

This margin left to “the person skilled in the art” can possibly be understood in certain relatively simple technical fields, to avoid circumventing a patent too easily. Let’s take the example of a patented process for driving a nail with a handled tool, which offers an object called a “*hammer*” as the preferred tool, which is also protected. Patent law may consider that the use of a tool other than the hammer – not described or claimed in the patent – is also covered by the patent insofar as it performs the same function as the hammer. An infringement judge applying the “*Doctrine of Equivalents*” will do the same [14].

But what about biotechnologies and in particular patents on genetic sequences? It can be very complex, if not impossible, for a “person skilled in the art” or an infringement judge to form an objective idea of the scope of a claim if the description of the invention in the patent application is not sufficient to support such a scope. The issue of patents on antibodies sums things up quite well. An antibody can be defined in the claims *via* the DNA sequence that encodes it or via its own sequence (it is a protein). But a much broader way of doing this is to adopt functional language, saying for example which receptor(s) it binds to. Regardless of its sequence. But it is easy to argue that such a definition is insufficient and uses terms that are too vague to justify such a broad scope. Moreover, case law in Europe and, very recently, the United States has calmed the ardour of applicants for such patents [15].

Hypothetical validity

Patents granted with this type of claim at the end of the application examination procedure may nevertheless be invalidated by the EPO's Boards of Appeal or national courts.

For example, a decision by a Board of Appeal of the EPO, concerning an application filed by the Max Planck Institute, established that a genetic sequence (nucleic acid or protein) whose structure is indicated in an application but whose function is indeterminate, obscure or only vaguely mentioned does not meet this condition. In the United States, in the "*Fisher*" decision [\[16\]](#), the Court of Appeal held that none of the uses claimed by the patent met the utility requirement because they represented only hypothetical possibilities and were clearly not "specific".

The role of a patent is to prevent third parties from infringing the protected invention. What happens if patents with broad and vague claims resist the litigation stage? Their owners will be the only ones to decide on the strategy to adopt, which is particularly delicate in certain fields and contexts. The field of plants, biodiversity and agriculture has its own specificities, particularly because it affects food, the environment and health. These aspects will be addressed in a later article.

Date Created

12 Jul 2023