Biotechnology Research and Innovation (2019) 3, 91–102

Biotechnology Research & Innovation

http://www.journals.elsevier.com/biotechnology-research-and-innovation/

RESEARCH PAPER

Sufficiency of disclosure and genus claims for protection of biological sequences: a comparative study among the patent offices in Brazil, Europe and the United States

Poliana Belisário Zorzal[a,b,d,*], **Fabricia Pires Pimenta**[b,d], **Antonio Alberto Ribeiro Fernandes**[a], **Alexandre Guimarães Vasconcellos**[c,d]

[a] *Universidade Federal do Espírito Santo (UFES), Núcleo de Biotecnologia, Vitória, ES, Brazil*
[b] *Fundação Oswaldo Cruz (FIOCRUZ), Centro de Desenvolvimento Tecnológico em Saúde, Rio de Janeiro, RJ, Brazil*
[c] *Instituto Nacional da Propriedade Industrial (INPI), Divisão de Pós-graduação e Pesquisa, Rio de Janeiro, RJ, Brazil*
[d] *National Institute of Science and Technology for Innovation on Diseases of Neglected Populations (INCT-IDPN), Rio de Janeiro, RJ, Brazil*

Received 25 September 2018; accepted 24 October 2018
Available online 10 November 2018

KEYWORDS
Biological sequences;
Sufficiency of disclosure;
Support;
Clarity;
Genus claims

Abstract Patent protection has been chosen as a strategy to protect new developments in molecular biology such as novel genes and proteins. A way to ensure the protection of genetic inventions is to claim a set of sequences that are associated with the described genetic sequences in terms of structure and/or biological activity, in a genus claim. Clearly, achieving an effective patent protection for proteins and genetic sequences is a real challenge for an Intellectual Property manager, considering the unpredictability of biological sciences and the diversity in current patent law and patent office guidance in each territory. This paper seeks to study the Brazilian patent office procedures about genus claims for biological sequences while comparing them with two other national/regional offices. To achieve this result, we initially present the concepts, followed by the current requirements and the barriers to obtain genus claims for biological sequences in the legal framework and patent office prosecution of Brazil, the European Union, and the United States. Subsequently, we study the impacts of these regulations in the scope of claim protection in each territory. This is done by comparing patent documents with the same priority granted in each of these offices in order to analyze the extension of the owner's rights for biological sequences. Understanding the logic that supports the examination procedures in the three studied offices will be important to subsidize the legal protection for gene-based inventions. Therefore, this would support the development of a patent system that can provide satisfactory safeguard for the results of investments in biotechnology Research and Development initiatives.

* Corresponding author.
E-mail: poliana.zorzal@cdts.fiocruz.br (P.B. Zorzal).

https://doi.org/10.1016/j.biori.2018.10.001

Two current ambitious initiatives, the Earth BioGenome Project (Lewin et al., 2018) and the Global Virome Project (Carroll et al., 2018), aim to sequence, catalogue, and characterize the genomes of all of Earth's eukaryotes and viruses over a period of 10 years in parallel initiatives. These projects highlight not only the increased necessity of scientific knowledge and understanding of Earth's biodiversity and virome, but also the importance of biological sequences for bioeconomy as the Fourth Industrial Revolution (Lewin et al., 2018). As result of this kind of research, scientists will be able to access locked biological and virome information necessary for maximizing returns to society and human welfare, providing timely data for public health interventions; new strategies for controlling outbreaks diseases; as well as the development of new products and processes.

In the pursuit of these ultimate goals, the biotechnology science and industry play an important role manipulating genes, genomes and metabolism through direct modification of DNA sequences (Carlson, 2016) in order to change the natural resources and generate an industrial utility for these new developments. All the R&D efforts are performed by public and private institutions planning future revenues with the development's negotiation and commercialization. In an increasingly knowledge-driven economy, the upcoming revenues are based in the patent system rights to exclude all competitors to explore economically the invention, thus assuring in a strategic market advantage to the patent developer.

On the one hand, the patent is a government-granted monopoly on an invention whereas, on the other hand, the patent document must meet some technical and legal requirements. In this context, one legal requirement especially important for biotechnological inventions is the sufficiency of disclosure. Non-natural DNA and amino acid sequences, when eligible to patentability in a specific territory, need to be described as clearly and fully as required by the patent office, possibly by the literal description and citation of every claimed biological sequence in the patent document.

Sufficiency of disclosure is a requirement responsible for the materialization of the principle of reciprocity in the patent system. Latin expressions as *Quid pro quo*[1] and *do ut des*[2] summarize the concept: wherein the government grants exclusive rights to the patent applicant in exchange for the full description of the invention teachings. These teachings will be available to third parties after the confidentiality phase of the patent application which is 18 months. Besides that, after twenty years of exclusive rights, will be freely available for the society make and use it in any lawful way.

[1] Quid pro quo (''something for something'' in Latin) is a phrase used in English to mean an exchange of goods or services, in which one transfer is contingent upon the other; ''a favour for a favour''.

[2] The formula do ut des (''I give that you might give'') expresses the reciprocity of exchange in relations.

Why is the sufficiency of disclosure a key concept for biological sequences?

The biological information included in the genetic code flows from a DNA sequence to an RNA transcript, which is usually translated into amino-acids sequences to fold in proteins. Each triplet in the nucleic acid sequence represents an amino-acid added to the protein primary structure. Considering the degeneracy of the genetic code and the existence of noncoding triplets, the 64 possible combinations of nucleotides encode only 20 different amino acids.

This feature of the genetic code allows that, in some cases, one can obtain the same amino acid sequence, and thus, the same protein from two or more different DNA or RNA sequences. Another characteristic of proteins is the possibility to behave identically even with significant differences in the amino acid sequence. In some proteins, more than 50% of the amino acids can be changed without substantially changing the protein function (Holman, 2004). Therefore, in order to assure the patent protection of genetic inventions, it has been customary to claim not only the genetic sequence literally described in the patent Description (the formal requirement) but also a set of related sequences (Dufresne & Duval, 2004). As a consequence, protection of biological sequences around the world has become a combination of a broad, heterogeneous and unstandardized claim strategies adopted by patent drafters (Yoo, Ramanathan, & Barcelon-Yanga, 2005) furthered by an unclear and excessively restricted examination procedure adopted by patent examiners. This landscape generates a high degree of uncertainty about the protection of sequences which demanded significant investments in terms of money, time and human resources to be developed.

One possible means to assure the protection of genetic inventions is to claim a set of sequences associated, even remotely, in terms of structure and/or biological activity, to the described genetic sequence (Cole, 2015). Therefore, patent drafters define a set of related sequences as a genus, particularly if there is reason to believe that the other sequences in the set have the same biological activities (McTavish, 2001). Strategies as Markush formulae, the percentage of identity, the percentage of similarity, hybridization with the specified sequence, homology (improperly percentage) and specific positions and types of substitutions are the main types of genus claims for a set of sequences (Table 1).

On the other hand, in some countries, including Brazil, it is usually issued a narrow protection that only covers the described biological sequence. Therefore, a competitor may minimally change the sequence and evade the patent, avoiding the infraction, while, at the same time, preserving the therapeutic efficacy of the encoded protein (Giles, 2011).

Considering the importance for the society of a fair scope of protection to the patent owner, it is important to discuss this issue in every national or regional patent office. Only in this way the patent system can perform its reciprocity through the sufficiency of the disclosure requirement.

As a globally controversial issue, the sufficiency of disclosure of genus claims for biological sequences has been an important topic appreciated in the patent office's procedures and legal disputes in some jurisdictions, sanctioning

Table 1 Examples of genus claims strategies used by patent drafters.

Type	Biological sequences genus claim	Example
Percentage of identity	All sequences have a threshold level of percentage of identity with the specified sequence	A phytase which has at least 74% identity to SEQ ID NO:2 and which comprises at least one alteration as compared to SEQ ID NO:2 in at least one position selected from the following: 52C/99C, 141C/199C, 59C/100C, 91 C/46C, 31C/176C, 31C/177C, and/or 162C/247C.
Markush formulae	A list of alternatively useable biological sequences recited in the claim.	A peptide or pseudo-peptide according to any of the preceding claims, characterized in that it corresponds to the general formula (I): XMPRY (I) wherein X is a group comprising 1–11 naturally occurring amino acids and/or nonconsecutive positive, Y is a group comprising 1–11 naturally occurring amino acids and/or nonconsecutive positive, X or Y contains at least one amino acid residue allowing the formation of a ring within the peptide, M denotes methionine or a isosteres or an analogue thereof, P denotes proline or a isosteres or an analogue thereof, R is arginine or its isosteres or an analogue thereof, and wherein the total number of amino acid residues is less than or equal to 25.
Hybridization	Capacity for hybridization with the specified sequence under certain experimental conditions	An isolated nucleic acid molecule that hybridizes under conditions of 2× SSC/0.1% SDS at 65°C. to said isolated nucleic acid according to claim 1.
Percentage of similarity	A threshold value for a percentage of similarity that they share with the literal sequence described	An isolated variant of a protein comprising the amino acid sequence shown in SEQ ID NO:3, wherein the variant comprises an amino acid sequence that is at least 95% similar to SEQ ID NO:3.
Function plus homology and percentage of homology	A threshold value for percentage of homology that sequences share with the specified sequence	An isolated enzyme with glucoamylase activity, which is connected with SEQ ID NO:7, and the degree of homology between the full-length sequences shown in 7 is at least 99%, and has an isoelectric point of less than 3.5 measured by isoelectric focusing.
Variation in specified position	A general mode that specifies the positions that vary (substitutions, deletions and additions) relative to the specified sequence	A phytase characterized by comprises at least one alteration and no more than 4 alterations as compared to SEQ ID NO:2 wherein at least one of said one to four alterations is selected from the following: N4P, N31C, W46E, K107G, Q111P, E119K, S162C, D202N, Q223E, E241Q, M273L, T276K, N286Q, I362K,R, I379K, N385D, G52C/A99C, G59C/F100C, Q111P/E241Q, K141C/V199C, S162C/S247C, N31C/T177C and W46C/Q91C, and wherein the phytase has an improved thermostability compared to SEQ ID NO:2.

Table 1 (*Continued*)

Type	Biological sequences genus claim	Example
Protein encoded by nucleic acid sequence	Amino acid sequence used to define the nucleic acid sequence	An isolated polypeptide having glucoamylase activity, selected from the group consisting of a polypeptide encoded by a polynucleotide having at least 65% sequence identity to the mature polypeptide coding sequence of SEQ ID NO:1 or SEQ ID NO:3, or the cDNA sequence thereof.
Nucleic acid encoding a protein	The nucleic acid encodes a polypeptide comprising the amino acid sequence as presented in the sequence listing in patent application	An isolated nucleic acid encoding a human Akt3 protein comprising the C-terminal sequence Cys-Gln-Gln-Ser-Asp-Cys-Gl Met-Leu-Gly-Asn-Trp-Lys-Lys, or a sequence in which more than 90% or 95% of the amino acids are identical to those of said sequence, wherein the nucleic acid encodes a polypeptide comprising the amino acid of SEQ ID NO:2.

Source: Author's elaboration.

for its allowance in some level. But the Brazilian patent office (Industrial Property Office – INPI) position is banning most categories of genus claims.

This paper seeks to study the Brazilian patent office procedures about genus claims for biological sequences compared with two other national/regional offices. To achieve this result, we present initially the concepts (Table 2), current requirements and the barriers to obtain genus claims for biological sequences in the legal framework and patent office prosecution of Brazil, European Union, and the United States. Subsequently, we study the impacts of these regulations in the scope of claim protection in each territory. It was done by comparing patent documents from the same priority granted in each of these offices in order to analyze the extension of the owner's rights for biological sequences.

This overview is an important step to provide a basis for a critical evaluation about the granted scope in each territory, as a reflex of the patent office interpretation for a legal concept. Moreover, in future perspective, this study can lead to the development of some level of harmonization and standards for protection of biological sequences in a global and/or regional level in order to strike a balance between the economic needs of industry and the public interest regarding the patent system.

Current domestic provisions about biological sequences genus claims

Brazil

According to the general Guidelines for Patent Applications Examination (INPI, 2013), generic disclosures in the Description would imply an extension of subject matter protection, resulting in a very strict acceptance of genus claims by INPI examiners. A broad claim, such as a genus claim, would be an object of irregularity (INPI, 2013), supposedly based on article 24 of the Industrial Property Law – Law 9.279/96 (Brasil, 1996). The article requires a complete description for the claimed invention, giving all information necessary to be carried out by a skilled person and, when possible, the best mode of execution.

Complementing the general guidelines, the Brazilian Guidelines for Patent Applications Examination in Biotechnology (INPI, 2015) and the patent prosecution by INPI give us the guidance that the specific biological sequences claimed in a species claim should always be presented literally in particular patent item, the ''Sequence Listing''. Accordingly, the presented sequences should be identified throughout the patent text by the term ''SEQ ID NO:'' (sequence identification number) followed by its corresponding numbering. Nucleotide and amino acid sequences claimed should be always characterized by their ''SEQ ID NO:'', the linear structure that unambiguously identifies the biological sequence.

According to INPI, in some cases, other forms of appearance or characterization of biological sequences may be accepted, such as: (a) When the sequences are smaller than four amino acids or ten nucleotides, the sequence itself must be in the Description (formal requirement) and not in the ''Sequence Listing''; (b) Structural formulas accompanied by their corresponding ''SEQ ID NO:''; (c) Markush formulas accompanied by their corresponding ''SEQ ID NO:''; (d) Deposit number; or (e) by its name or designation if the biological sequence is already known in the prior art and is not the principal object of the invention.

In a genus claim, the protection beyond the sequence what was developed by the patent owner would expand the patent exclusivity for more than would be possible to be performed by a skilled person after the monopoly period. As a consequence, the society will receive a narrower scope of knowledge and teachings when compared to what was claimed by the patent owner and monopolized for 20 years.

Given this scenario, genus claims are frequently interpreted as broader than the invention disclosed, if covering

Table 2 The terminology on sufficiency of disclosure adopted in tree patent offices and their respective legal basis and applicable rule: INPI (Brazilian Industrial Property Institute), EPO (European Patent Office), and USPTO (United States Patent and Trademark Office).

	INPI		EPO		USPTO	
	Legal basis	In practice rule	Legal basis	In practice rule	Legal basis	In practice rule
	Industrial Property Brazilian Law (9.279/96)	Guidelines for Patent Applications Examination	The European Patent Convention	Guidelines for Examination	35 U.S. Code	Manual of Patent Examining Procedure
Disclose the invention in a manner sufficient clear and complete for it to be carried out by a person skilled in the art	Art. 24	Suficiência descritiva/sufficient description (2.13)	Art. 83	Sufficiency of disclosure (Part F, Chapter III, 1.)	§ 112 First Paragraph	Enablement (MPEP2164)
Claims shall be supported by the description defining the matter for which protection is sought	Art. 25	Fundamentação/substantiation (3.85)	Art. 84	Support in description (Part F Chapter IV, 6.)	§ 112 First Paragraph	Written description (MPEP2163)
The claims shall be clear and precise	Art. 25	Clareza/clarity (3.36)	Art. 84	Clarity (Part F, Chapter IV, 4.)	§ 112 First Paragraph	Definiteness (MPEP 2173)

Source: Author's elaboration.

many different sequences of amino acids or nucleotides without specifying the locations for replacements. In these cases, the Description would not provide sufficient information to allow a skilled person to make and use all the claimed possibilities of the invention since these claims cover a huge number of undisclosed biological sequences. For a description to be considered sufficient, the patent applicant must present all sequences covered by a genus claim, according to INPI guidance.

Regarding the requirement of support in the description and clarity of claims, biological sequences claimed in a genus based on a generic characterization is considered by INPI unsupported as well as unclear and imprecise. Article 25 of the LPI states that ''claims must be supported on the Description, characterizing the particularities of the application and defining clearly and precisely the protected subject matter''. Thus, according to INPI guidance, a broad claim including a set of sequences without their description would not be supported by the Description. In this sense, ''protein characterized by consisting in the amino acid sequence encoded by the nucleic sequence SEQ ID NO: 2'', ''DNA sequence characterized by codifying a polypeptide'' and ''% of identity'' are claims objected by INPI based in the argument of clarity lack (INPI, 2015).

The presentation as Markush formula is the unique option given by INPI guidance for biological sequence genus claims. Through Markush formulae is possible to present a basic sequence definition for amino acids or nucleotides and to provide alternatives of variable or optional units in different positions, accompanied by a list of definitions of said groups. As a result, a plurality of sequences may be protected from a single formula representation. This formula holds in particular whenever a broad genus and/or functional features must be protected (Tostmann, 2015), as in general the variable units have similar properties, as physical–chemical features.

However, the Markush formula is not completely suitable as an alternative to biological sequences claims, since it is necessary that all possible representative substituents in the claimed Markush formula must be supported on the Description, clearly and precisely defined. Therefore, in Brazil, providing a Markush formula could be an option as a genus claim accepted by INPI, but will not solve the problem of narrow claims about biological sequences.

European Union

In the European Patent Convention (EPC), the sufficiency of disclosure requirement is defined in article 83, which states that ''the European patent application shall disclose the invention in a manner sufficiently clear and complete for it to be carried out by a person skilled in the art''. Article 84 states that ''the claims shall define the matter for which protection is sought. They shall be clear and concise and be supported by the description''.

For an appropriate description of claimed biological sequences, a unique example could be enough. But in the case of genus claims for biological sequences, the description must provide a number of examples or describe alternative embodiments in order to meet the sufficiency of description requirement (Latimer, 2005). Besides examples, the patent description must give all necessary information to carry out all the claimed scope, without undue experimentation or inventive step (T 0727/95).

According to the European Patent Office (EPO)'s Implementing Regulations, in some technical fields, a clear description of function may be much more appropriate than an exhaustive description of the structure (EPO, 2015). In this context, the sufficiency of disclosure for biological sequences could be fulfilled if the invention also describes a biological function.

For the EPO, the applicant can claim all obvious modifications of, equivalents to and uses of the sequence which he has described in the application (EPO, 2017). This is essentially equivalent to a genus claim. Further, if it is reasonable to expect that all the variants covered by a genus claim have the properties or uses assigned in the disclosure, the applicant encouraged to seek full protection (T 0172/99, T 1727/12, and Biogen Inc. v Medeva plc). According to the EPO's Guidelines for Examination (EPO, 2017), a genus claim may be acceptable, even in a broad scope, where there is reasoning in the Description and there is no reason to suppose that the invention cannot be extended across the entire claimed category.

This approach is related to an EPO's interpretation that a large number of claims are some generalization of one or more particular examples. However, the office assumes that the permissible degree of generalization is a matter which the examiner must judge in each particular case in the light of the relevant prior art. In this sense, a fair claim under the EPO approach would be that which is not so broad as to extrapolate what was developed by the applicant and is not so limited as to deprive the applicant of a fair monopoly by disclosing the invention.

In line with the EPO's issue treatment, the European courts in the face of many internal controversial decisions have sought to focus on an interpretation considering mainly the purpose of the sufficiency of disclosure in the patent system. Examples as the percentage of identity claims could be used by a patent drafter but disclosing the fragments where the percentage is applied (T 1644/08). However, it is not necessary to define how this identity should be calculated in the sequence as the methods to calculate the degree of identity among different sequences were known to the skilled person (T 1644/08). If hybridization sequences are claimed, the exact conditions of hybridization and the hybridizing fragment length should be clearly indicated in that claim, providing details about the degree of genetic similarity between the sets of DNA sequences (T 0837/07).

In this scenario, it is clear that the European approach to genus claims for biological sequences is supported by a generic and/or narrower disclosure description focused on sequence functional knowledge, specification on the way of comparison, and not only on linear structure. Additionally, the European system allows functional language and generally requires fewer species to be described in order to enable a broader genus claim. This more balanced vision established an evolved patent system to the biotechnological sector in Europe.

United States

According to United States Patent Act § 112 (OLRC, 2012) the patent Specification (formal requirement) shall contain ''a written description of the invention, and of the manner and process of making and using it, in such full, clear, concise, and exact terms as to enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make and use the same, and shall set forth the best mode contemplated by the inventor or joint inventor of carrying out the invention''. The ''written description requirement'' is interpreted as encompassing three separated requirements: written description, enablement and best mode, where the first two are especially important to biological sequence inventions.

In order to meet the ''written description requirement'', the Specification of a patent should clearly state what was invented according to *In re* Barker (1977) and cover the full scope claimed in the patent application. The description of the subject matter actually invented shows that the patentee was in possession of the invention when the patent was filed in a land rights case (Capon v. Eshhar, 2005).

In the case of biological sequence claims, the United States Patent and Trademark Office (USPTO) and courts have concurred with the description of a representative number of sequences within the scope of the genus claim or citing common structural features to a substantial portion of the genus, demonstrating the sequence possession. This interpretation is based in the pre-existing knowledge about genetic code redundancy and bioinformatics data, which could be used for identifying all possible sequences within a certain genus (USPTO, 2008).

The categories created by a genus claim can encompass sequences which match only partially with the disclosed amino acid or nucleotide sequence, allowing variations in the described sequence. In this context, the USPTO guidance is the inclusion of additional information regarding which of the amino acids or nucleotides may vary in comparison with the presented sequence.

Knowledge about protein activity suggesting that similarity of structure confers the claimed activity can be essential information for a genus claim allowance. In USPTO interpretation, for amino acid sequences a known or disclosed correlation between a structure other than the disclosed one and a claimed activity can be enough to allow one of skill in the art to identify other proteins expected to have the same or similar tertiary structure (USPTO, 2009). Accordingly, in specific cases, one skilled in the art would accept the disclosure of a sequence as representative of other proteins having a claimed activity, but this representative number can vary greatly (Abbvie v. Janssen, 2014).

To summarize the USPTO and the US courts interpretations, the patent law requires a specific description of the acceptable sequence variations, clearly more than merely mentioning a range of identity or similarity of a referenced sequence, claiming it and its functional variants. This specific description includes a sequence, structures, biological features, physical and chemical properties and also that an inventor cannot claim patent rights on a DNA invention before its actual creation.

In addition, to meet the written description requirement, the claims of a patent application must be enabled by explaining how to make and use the invention without ''undue experimentation''. In other words, the patent specification must teach one ''skilled in the art'' how to practice the invention and this amount of ''teaching'' required depends on the invention (Kellam, 2001). This requirement assures that the public will be able to use the technology when the patent expires (United States v. Dubilier Condenser Corp., 1933), allowing one of ordinary skill in the art to practice the invention.

Considering the enablement requirement, predictability is also a relevant factor to be considered for determining enablement. Some knowledge areas, including Biotechnology, are not considered predictable since a single embodiment of the invention will not provide a broad enablement. For predictable areas, such as mechanical arts, for instance, enablement is much easier and can be based on the Specification and on known scientific laws (*In re*Fisher, 1970). However, in general, biotechnological areas, the researchers are not able to predict, for example, how a change in the amino acid sequence will affect a protein function (Sampson, 2000) in a broad sense. Since unpredictability limits the breadth of the claims, Biotechnology arts, in general, have narrower scopes than other areas.

During the enablement requirement analysis, the examiner considers the working examples, which, in unpredictable areas, are an important factor for enablement. The USPTO assumes that representative examples together with an applicable disclosure of the whole genus are sufficient for enablement without undue experimentation. However, in Biotechnology, proofs of enablement may be required for other members of the claimed genus, according to the invention predictability.

Also, the US claims must be definite, i.e. must be written so that a skilled in the art would understand to where the boundaries of the patent right extend. Definiteness requirement is analyzed in light of the content of the particular application disclosure, the teachings of the prior art, and the claim interpretation that would be given by one possessing the ordinary level of skill in the pertinent art at the time the invention was made.

Comparative study on real cases

As we have seen, there are strong differences amongst jurisdictions concerning what disclosure is required and what claim breath is allowed for biological sequences. In Brazil, the INPI interpretation about the sufficiency of disclosure for genus claim of biological sequences has been very strict. In this scenario, we present a comparison among correspondents of a patent family regarding a genus claim for biological sequences and patent office's examination procedures to identify the results of each of the national office procedures.

We shall examine three patents of a same family by INPI, EPO and USPTO to understand the scope of granted claims. The following patent documents were compared, especially regarding to sufficiency of disclosure requirements for genus claims of biological sequences:

- INPI - PI0709732-8
- EPO - EP2365064
- USPTO - US 8,460,656

The compared invention relates to a phytase which has identity in the amino acid sequence to a phytase derived from *Citrobacter braakii* (the mature part of the *Citrobacter braaki* ATCC 51113 phytase is included in the sequence listing as ''SEQ ID NO:2'') and comprises alterations when compared to this phytase. These phytase variants have improved properties, such as thermostability, temperature profile, pH profile, specific activity, among others. The patent documents also relate to DNA sequences encoding these phytases. A comparison of the first independent claim as issued and as granted in Brazil, Europe, and the US is given in Table 3.

Originally, claim 1 presented by the applicant in INPI and USPTO is the same, referring to an isolated polypeptide presenting phytase activity having an identity of 74% with a specific sequence provided with at least one amino acid altered in a specified position. In EPO the patent claim is narrower, as the number of possibilities for eventual alterations given is much less (92 alterations in INPI and USPTO and 7 in EPO).

In the granted claim of the Brazilian document, the percentage of identity categorization is absent, while the proposed percentage of 74% was changed to 85% in EPO and to 80% in USPTO, stricter numbers for identity with the sequence IDs provided in the claims.

An important point is the inclusion of a restriction related to functional improvement in the Brazilian and European granted claim, while in the US there was no limitation for an improved thermostability compared to the phytase of SEQ ID NO:2. As a result, the final American claim covers a range of phytases with no clear indication of increased effects/utility (Ravi, 2013) when compared to the prior art, extending the scope of protection.

In order to guarantee a balanced scope of protection, even in the Brazilian granted claims in which the improved activity is claimed, an objective standard about this improvement could be more suitable. A specific number, emphasizing objectively and quantitatively the claimed improvements and indicating the level of optimization in protein thermostability, supported by the Description and Examples, should be part of a genus claim. Such quantification can be observed in the European claim ''where the residual activity of the phytase is at least 120% of the residual activity of the reference phytase SEQ ID NO:2 measured in the same conditions, and which comprises at least one alteration as compared to SEQ ID NO:2 in at least one position selected from the following (...)''. This quantification guarantees the protection of a range of sequences with the minimum thermostability's improvement claimed by the granted claim, but it is absent about an upper limit of improvement, balanced and related to the invented phytase. In this case, even the European claim overprotect sequences with better thermostability than the developed by the patent owner.

Another issue with regard to complete specification is related to the specified conditions described in the European granted claim for the improvement of protein thermostability. For each kind of genus claim, it is crucial to define the conditions under which the sequence comparison should be done.

For example, when using the identity percentage approach to claim biological equivalents or analogues, it is critical to define explicitly the scoring matrix and the gap penalties, either in the claims themselves or in the definition of similarity or identity percentage provided in the Description. In the European claim, we can observe this definition of the percentage of identity ''wherein the degree of identity between sequences, is determined by the programme ''Needle'' using the substitution matrix BLOSUM62, the gap opening penalty is 10, and the gap extension penalty is 0.5''.

Another example of parameter description is when using the hybridization language genus claim approach which must contain the chemical and physical conditions for hybridization occurrence, specifically, the concentrations and temperatures required. These ''stringency'' factors must be defined in the Description and/or in the claims in order to determine conditions to limit the exact claim boundaries. It is clear for a person skilled in the art that the level of stringency that delimits the boundary of the set of sequences claimed is extremely variable from one patent to another, so it is very important a clear definition of the parameters.

There is no doubt that among all evaluated patent offices, the stricter claim and narrower protection is given by the Brazilian one, where the claim protection is related to specified alterations in the disclosed sequence.

Discussion and conclusions

Patent protection is in the core of life sciences and biotechnology businesses and these fields require a reasonable breadth of protection, considering their peculiarities. This is especially true when biological sequence claims are considered, and the unpredictability of these arts becomes a barrier for some claim generalizations allowed in other technical areas. This biotech unpredictability can lead to excessively narrow protection for biological sequences and denial of genus claims, which is assumed in many cases in order to guarantee a reasonable monopoly to the patent's owner.

Especially in the Brazilian context, the examiner's everyday practices and the office's guidance poses additional challenges and limits to the sufficiency of disclosure legal requirements for biological sequences. But, imposing strict criteria on genus claims for biological sequences, i.e., only allowing claims that encompass specific examples or disclosure in the description, can generate difficulties to enforce a patent. A limited scope for patent enforcement, when compared to other patent offices as we found in this study, could allow the competitor's design around the granted sequence by replacing a few irrelevant nucleotides or amino acids. It could be a discouragement for biotechnological developments in the Brazilian context.

The sufficiency of disclosure legal requirement, as a crucial driving force in which patent system is based, allows the Brazilian patent system to take into account the special characteristics of biological sequences in order to accept the notion of a genus including, but not restricted to, the sequence analogues. Therefore, as a viable and legitimate

Table 3 Comparison of the first independent claim as filed and granted by Brazilian, European and American National Patent Offices.

	Claim as filed	Claim as grant
INPI	1. A phytase which has at least 74% identity to SEQ ID NO:2 and which comprises at least one alteration as compared to SEQ ID NO:2 in at least one position selected from the following: 1, 2, 3, 4, 5, 31, 41, 46, 52, 53, 55, 57, 59, 74, 76, 82, 84, 91, 99, 100, 104, 105, 107, 109, 111, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 136, 137, 141, 154, 161, 162, 164, 167, 171, 176, 177, 179, 180, 181, 182, 183, 184, 185, 186, 196, 199, 200, 202, 203, 218, 223, 239, 240, 241, 247, 273, 276, 281, 282, 283, 284, 285, 286, 289, 294, 299, 308, 314, 316, 324, 331, 339, 351, 355, 362, 379, 385, 406, 409, 410, and 411; with the proviso that the phytase is not SEQ ID NO:3, not SEQ ID NO:4, and not SEQ ID NO:6.	1. A phytase characterized by comprises at least one alteration and no more than 4 alterations as compared to SEQ ID NO:2 wherein at least one of said one to four alterations is selected from the following: N4P, N31C, W46E, K107G, Q111P, E119K, S162C, D202N, Q223E, E241Q, M273L, T276K, N286Q, I362K,R, I379K, N385D, G52C/A99C, G59C/F100C, Q111P/E241Q, K141C/V199C, S162C/S247C, N31C/T177C and W46C/Q91C, and wherein the phytase has an improved thermostability compared to SEQ ID NO:2.
EPO	1. A phytase which has at least 74% identity to SEQ ID NO:2 and which comprises at least one alteration as compared to SEQ ID NO:2 in at least one position selected from the following: 52C/99C, 141C/199C, 59C/100C, 91 C/46C, 31C/176C, 31C/177C, and/or 162C/247C.	1. A phytase which has at least 85% identity to SEQ ID NO:2, which has an improved thermostability indicated as residual activity determined by dividing a fermentation supernatant in two parts, one part is incubated for 30 min at 60 °C, and the other part for 30 min at 5 °C, following which the activity of both is determined on p-nitrophenyl phosphate at 37 °C and pH 5.5, where the residual activity of the phytase is the activity of the sample having been incubated at 60 °C divided by the activity of the same sample having been incubated at 5 °C, where the residual activity of the phytase is at least 120% of the residual activity of the reference phytase SEQ ID NO:2, measured in the same conditions, and which comprises at least one alteration as compared to SEQ ID NO:2 in at least one position selected from the following: G52C/A99C, K141C/V199C, G59C/F100C, Q91C/W46C, N31C/E176C, N31C/T177C, and/or S162C/S247C, wherein the degree of identity between sequences, is determined by the programme ''Needle'' using the substitution matrix BLOSUM62, the gap opening penalty is 10, and the gap extension penalty is 0.5.
USPTO	1. A phytase which has at least 74% identity to SEQ ID NO:2 and which comprises at least one alteration as compared to SEQ ID NO:2 in at least one position selected from the following: 1, 2, 3, 4, 5, 31, 41, 46, 52, 53, 55, 57, 59, 74, 76, 82, 84, 91, 99, 100, 104, 105, 107, 109, 111, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 136, 137, 141, 154, 161, 162, 164, 167, 171, 176, 177, 179, 180, 181, 182, 183, 184, 185, 186, 196, 199, 200, 202, 203, 218, 223, 239, 240, 241, 247, 273, 276, 281, 282, 283, 284, 285, 286, 289, 294, 299, 308, 314, 316, 324, 331, 339, 351, 355, 362, 379, 385, 406, 409, 410, and 411; with the proviso that the phytase is not SEQ ID NO:3, not SEQ ID NO:4, and not SEQ ID NO:6.	1. A phytase variant, comprising an amino acid substitution corresponding to an amino acid substitution in SEQ ID NO:2 selected from the group consisting of 4P, 31C, 31C/176C, 31C/177C, 46C/91C, 46E, 52C/99C, 59C/100C, 107G, 111P, 111P/241Q, 119K, 141C/199C, 162C, 162C/247C, 202N, 223E, 241Q, 273L, 276K, 286Q, 362K, R, 379K, and 385D wherein the variant has at least 80% identity to SEQ ID NO:2 and has phytase activity.

Source: Author's elaboration.

right, genus claims for biological sequences must be considered in examiners daily practices, with some important precautions.

One is the establishment of limits for each kind of genus claims, always defining the sequence genus in a positive way. When based on a sequence percentage of identity or similarity, the means to establish this percentage should be explicit in the patent office guidance, considering that different methods may provide different results. If a claim is not clear about the categorization method, the claimed categories can include different sequences inside, even using the same percentage number. Also, the hybridization claims under certain experimental conditions must be clearly defined in the official recommendation, as well as experimental conditions and level of stringency delimiting the set of sequences claimed. Furthermore, functional limitations for the sequences belonging to the claimed genus, including to emphasize objectively and quantitatively the claimed improvements, are an important strategy to restrict and to include in the protected genus only the functional equivalents and analogues.

A rational allowance of genus claims for biological sequences and standardized alternatives for sequence generalization should be established by the patent office, protecting biologically equivalents and analogues sequences to the literal ''SEQ ID NO:'' presented. In this regard, the formulation of clear rules covering submission and enforcement of patents of genetic sequence protected by genus claims would provide guidelines to be followed by the patent practitioner and at the end of the chain, to stimulate investments in research and innovation in biotechnology in Brazil.

Anyway, supplementary to genus claim acceptance by the patent office, strategies adopted by the applicants for drafting patent applications for biological sequences should include as many as possible ways to define the pursued sequences, including physiochemical properties of the polypeptide or protein and the procedure for processing the gene, polypeptide or protein. Moreover, since biotechnology is an experimental science, when drafting the patent Description of an invention, the applicant should describe the state of the art (such as the research and development of related variants and gene evolution) and provide as many as possible actual examples (Lung Tin, 2017). Such strategy will contribute to reaching the required level of sequence description for which a patent is sought in the Description (the concepts of fundamentação for INPI, support in description for EPO and written description for USPTO).

In addition, to assure a complete description for the claimed invention to be carried out, a helpful strategy when drafting the patent document is to describe the origin of the sequence, the technical means for getting it, and their functions and technical effects, domains and other key sections for the sequence. Such draft approach gives technical subsidies to describe the expected functions and effects of ''SEQ ID NO:'''s analogues and equivalents, based on the provided evolutionary information and experimental data between different strains or species, targeting a supported broader protection.

Achieving a fair and effective patent protection for proteins and genetic sequences is a real challenge, considering the difficulty to establish standards for nucleotides and protein structure and function, the existence of analogues and biological equivalents and the unpredictability of biological systems. By the same token, it is hard for patent drafters and owners to take into account strong differences amongst jurisdictions concerning the requirements for sufficiency disclosure in a patent application.

Although some level of extension in protection is desirable, broader than to the literally described biological sequences, a clear limitation is required for the protection of sequences included in a genus. In the long run, cases where the genus claim is granted without additional and clear limitation, as in the US claim mentioned in this study, can hinder innovation and development in the subject matter (Ravi, 2013). One must not forget that the ultimate goal of patent protection is to promote innovation for the benefit of all society.

The right balance between compensation and incentivizing improvements and competition may be particularly difficult for in the biotechnology field, which a number of similar yet chemically distinct sequences may perform equivalent functions (De Luca & Trifonova, 2017). The present study gives us a realistic scenario of this pronounced duality: on one side is the patent owner, seeking broad protection – ideally, not broader than the invention, but broader enough to stop competitors from using innocuous alterations of the patented sequence, whose development entailed vast amounts of capital (including for research and regulatory approval). On the other side, the official patent office acts seeking for granting a patent that reveals what a person skilled in the art will be able to perform, after the end of the given monopoly, based in what was effectively disclosed. It allows a balance between inventor and society's interests.

Exclusivity rights based on patent documents failing to provide desirable public disclosure for biological sequences or granting a narrower protection for the sequences than those possibly performed by a person skilled in the art are unfair, as well as innovation inhibitors (Zhang, Sherwinter, & Greenbaum, 2017). Given the importance of the biotechnology solutions to the society, it is expected that patent offices and courts will continue to refine the boundaries of what is required for claims disclosure encompassing a number of biological equivalents, towards a balanced protection and some standardization in the patent drafting.

The evolution of gene-based inventions has historically led to higher standards for specification in the legal frameworks and patent prosecution in the patent offices investigated. But, as computational search and comparison tools for genetic sequences are evolving rapidly, it is becoming reachable to determine the a range of genetic sequence possibilities that can result in a particular biological function. Our study shows that, even as an unstandardized process in and between the studied jurisdictions, the genus claims have been assumed as an inevitable tendency able to fulfil the sufficiency of disclosure for biological sequences.

Conflicts of interest

The authors declare no conflicts of interest.

Funding

This study was financed in part by the Coordination for the Improvement of Higher Level Personnel (CAPES – Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) – Finance Code 001 and Oswaldo Cruz Foundation (FIOCRUZ).

References

AbbVie Deutschland GmbH v. Janssen Biotechnology, Ltd., 759 F.3d 1285. (Fed. Cir. 2014).

Brasil. (1996). Law n. 9.279 of May 14, 1996. Law on Industrial Property – Regulates rights and obligations related to industrial property. Accessed from http://www.wipo.int/wipolex/en/text.jsp?file_id=125397.

Biogen Inc v Medeva plc. (1997). RPC 1, [1996] UKHL 18, (1997) 38 BMLR 149, on appeal from the Court of Appeal – Biogen Inc v Medeva Plc [1995] FSR 4; [1995] RPC 25, and the Patents Court – Biogen Inc v Medeva Plc. (Unreported, November 4, 1993).

Capon v. Eshhar, 418 F.3d 1349, 1357, 76 USPQ2d 1078, 1084. (Fed. Cir. 2005).

Carlson, R. (2016). Estimating the biotech sector's contribution to the US economy. *Nature Biotechnology*, *34*(3), 247. http://dx.doi.org/10.1038/nbt.3491

Carroll, D., Daszak, P., Wolfe, N. D., Gao, G. F., Morel, C. M., Morzaria, S., et al. (2018). The global virome project. *Science*, *359*(6378), 872–874. http://dx.doi.org/10.1126/science.aap7463

Cole, P. (2015). Patentability of genes: A European Union perspective. *Cold Spring Harbor Perspectives in Medicine*, *5*(5) http://dx.doi.org/10.1101/cshperspect.a020891

De Luca, C., & Trifonova, A. (2017). Patent disclosure requirements for therapeutic antibody patents. *Expert Opinion on Therapeutic Patents*, *27*(8), 867–875. http://dx.doi.org/10.1080/13543776.2017.1296950

Dufresne, G., & Duval, M. (2004). Genetic sequences: how are they patented? *Nature Biotechnology*, *22*, 231–232. http://dx.doi.org/10.1038/nbt0204-231

EPO, European Patent Office. (2015). European Patent Convention. Implementing Regulations - to the Convention on the Grant of European Patents. Accessed from https://www.epo.org/law-practice/legal-texts/html/epc/2016/e/ma2.html.

EPO, European Patent Office. (2017). Guidelines for Examination in the European Patent Office. Accessed from https://www.epo.org/law-practice/legal-texts/guidelines.html. Fisher, 427 F.2d 833, 839. (C.C.P.A. 1970).

Giles, P. B. (2011). How to claim a gene: Application of the patent disclosure requirements to genetic sequences. *Georgia State University Law Review*, *27*(3), 6. https://readingroom.law.gsu.edu/gsulr/vol27/iss3/6

Holman, C. M. (2004). Protein similarity score: A simplified version of the blast score as a superior alternative to percent identity for claiming genuses of related protein sequences. *Santa Clara Computer & High Technology Law Journal*, *21*(55). https://ssrn.com/abstract=932070

In re Barker, 559 F.2d 588, 592 n.4, 194 USPQ 470, 473 n.4. (CCPA 1977).

In re Fisher, 427 F.2d 833, 839, 166 USPQ 18, 24 (CCPA 1970).

INPI, Brazilian Industrial Property Institute. (2013). Guidelines for Patent Applications Examination – Patent Application: Title, Description, Claims, Drawings and Summary. Resolution n. 124/2013. Accessed from http://www.inpi.gov.br/legislacao-arquivo/docs/resolucao_124_diretrizes_bloco_1_versao_final_03_12_2013_0.pdf.

INPI, Brazilian Industrial Property Institute. (2015). Brazilian Guidelines for Patent Applications Examination in Biotechnology. Resolution n. 144/2015. Accessed from http://www.inpi.gov.br/sobre/arquivos/resolucao_144-2015_-_diretrizes_biotecnologia.pdf.

Kellam, M. D. (2001). Making sense out of antisense: The enablement requirement in biotechnology after Enzo Biochem v. Calgene. *Indian Law Journal*, *76*, 221.

Lewin, H. A., Robinson, G. E., Kress, W. J., Baker, W. J., Coddington, J., Crandall, K. A., et al. (2018). Earth BioGenome Project: Sequencing life for the future of life. *Proceedings of the National Academy of Sciences*, *115*(17), 4325–4333. http://dx.doi.org/10.1073/pnas.1720115115

Latimer, M. T. (2005). Patenting inventions arising from biological research. *Genome Biology*, *6*(1), 203.

Lung Tin IP (2017). Newsletter – China's Recent Development on Granting Biological Sequence Claims. http://www.lungtin.com/UploadFile/Files/2017/12/7/151833146ae51e203-2.pdf. Accessed 31 July 2018.

McTavish, H. (2001). Enabling genus patent claims to DNA. *Minnesota Intellectual Property Review*, *2*, i. http://hdl.handle.net/10822/517537

OLRC, Office of the Law Revision Code. (2012). Title 35, United States Code, part II, chapter 11, section 112. Accessed from http://uscode.house.gov/view.xhtml?req=granuleid:USC-prelim-title35-section112&num=0&edition=prelim.

Ravi, B. (2013). Gene Patents in India: Gauging policy by an analysis of the grants made by the Indian Patent Office. *Journal of Intellectual Property Rights*, *18*, 323–329. http://nopr.niscair.res.in/handle/123456789/20283

Sampson, M. (2000). The evolution of the enablement and written description requirements under 35 USC § 112 in the area of biotechnology. *Berkeley Technology Law Journal*, 1233–1274. https://www.jstor.org/stable/24115688

T 0727/95 (Cellulose) of 21.5.1999. Decision of Technical Board of Appeal. Accessed from https://www.epo.org/law-practice/case-law-appeals/recent/t950727ep1.html.

T 1644/08 (Endoglucanases/NOVOZYMES) of 15.2.2011. Boards of Appeal of the European Patent Office. Accessed from https://www.epo.org/law-practice/case-law-appeals/recent/t081644eu1.html#q1644%2F08%20.

T 1727/12 of of 1.2.2016. Boards of Appeal of the European Patent Office. Accessed from https://www.epo.org/law-practice/case-law-appeals/recent/t121727eu1.html.

T 0172/99 of 7.3.2002. Boards of Appeal of the European Patent Office. Accessed from https://www.epo.org/law-practice/case-law-appeals/recent/t990172eu1.html.

Tostmann, H. (2015). Protecting chemistry inventions: The double-edged sword of being an unpredictable art. *ACS Medical Chemical Letters*, *6*(4), 364–366. http://dx.doi.org/10.1021/acsmedchemlett.5b00116

United States v. Dubilier Condenser Corp., 289 U.S. 178, 186-187 (1933).

USPTO, United States Patent and Trademark Office. (2008). U.S. Dept of Commerce, Written Description Training Materials. Accessed from http://www.uspto.gov/sites/default/files/web/menu/written.pdf.

USPTO, United States Patent and Trademark Office. (2009). Training materials for examining patent applications with respect to 35 u.s.c. section 112, first paragraph-enablement of chemical/biotechnical applications.

Yoo, H., Ramanathan, C., & Barcelon-Yang, C. (2005). Intellectual property management of biosequence information from a patent searching perspective. *World Patent Information*, *27*(3), 203–211. http://dx.doi.org/10.1016/j.wpi.2005.02.001

Zhang, T., Sherwinter, D. J., & Greenbaum, D. (2017). Call for Standardization in Patent Claim Drafting. *Santa Clara Computer & High Technology Law Journal*, *34*, 290. https://digitalcommons.law.scu.edu/chtlj/vol34/iss3/2