INSTITUTO NACIONAL DA PROPRIEDADE INDUSTRIAL

Revista da Propriedade Industrial

Nº 2608
29 de Dezembro de 2020

Comunicados

Seção I

REPÚBLICA FEDERATIVA DO BRASIL

Presidente

Jair Bolsonaro

MINISTÉRIO DA ECONOMIA

Ministro da Economia

Paulo Roberto Nunes Guedes

INSTITUTO NACIONAL DA PROPRIEDADE INDUSTRIAL

Presidente

Claudio Vilar Furtado

De conformidade com a Lei nº 5.648 de 11 de dezembro de 1970, esta é a publicação oficial do Instituto Nacional da Propriedade Industrial, órgão vinculado ao Ministério da Economia, República Federativa do Brasil, que publica todos os seus atos, despachos e decisões relativos ao sistema de propriedade industrial no Brasil, compreendendo Marcas e Patentes, bem como os referentes a contratos de Transferência de Tecnologia e assuntos correlatos, além dos que dizem respeito ao registro de programas de computador como direito autoral.

As established by Law nº 5.648 of december 11, 1970, this is the official publication of the National Institute of Industrial Property, an office under the Ministry of Economy, Federative Republic of Brazil, which publishes all its official acts, orders and decisions regarding the industrial property system in Brazil, comprising Trademarks and Patents, as well as those refering to Technology Transfer agreements and related matters, besides those regarding software registering as copyright.

D´après la Loi nº 5.648 du 11 décembre 1970, celle-si est la publication officielle de l'Institut National de la Propriété Industrielle, un office lié au Ministère de l'Économie, République Fédérative du Brésil, qui publie tous ses actes, ordres et décisions concernant le système de la propriété industrielle au Brésil, y compris marques et brevets, aussi que ceux référents aux contracts de transfert de technologie et des sujets afférents, en outre que ceux se rapportant à l'enregistrement des programmes d´ordinateur comme droit d'auteur.

Según estabelece la Ley nº 5.648 de 11 diciembre 1970, esta es la publicación oficial del Instituto Nacional de la Propiedad Industrial, oficina vinculada al Ministerio de la Economía, República Federativa del Brasil, que publica todos sus actos, ordenes y decisiónes referentes al sistema de propiedad industrial en Brasil, comprendendo marcas y patentes así que los referentes a contractos de transferencia de tecnologia y asuntos corelacionados, además de los referentes al registro de programas de ordenador como derecho de autor.

Laut Gezets Nr. 5.648 vom 11. dezember 1970, ist dies das Amtsblatt des Nationalen Instituts für gewerbliches Eigentum (INPI), eines Organs des Bundesministerium für Wirtschaft, der Bundesrepublik Brasilien, welches alle Amtshandlungen, Beschlüsse und Entscheidungen über gewerbliches Eigentum in Brasilien, einschliesslich Warenzeichen und Patente, ebenso wie auch Übertragunsvertrage von Technologie und Computerprogramme als Urheberrecht veroffentlicht.

MINISTÉRIO DA ECONOMIA
INSTITUTO NACIONAL DA PROPRIEDADE INDUSTRIAL

PORTARIA/INPI/PR Nº 405, DE 21 DE DEZEMBRO DE 2020.

O PRESIDENTE e a **DIRETORA DE PATENTES, PROGRAMAS DE COMPUTADOR E TOPOGRAFIAS DE CIRCUITOS INTEGRADOS** do **INSTITUTO NACIONAL DA PROPRIEDADE INDUSTRIAL**, no uso de suas atribuições previstas nos artigos 17, inciso XI, e 19 da Estrutura Regimental do INPI, aprovada pelo Decreto nº 8.854, de 22 de setembro de 2016, e inciso XII do artigo 152 do Regimento Interno, aprovado pela Portaria MDIC nº 11, de 27 de janeiro de 2017,

RESOLVEM:

Art. 1º Esta Portaria dispõe sobre os procedimentos para a apresentação da Listagem de Sequências, para fins de complementação do relatório descritivo constante dos pedidos de patentes depositados no INPI a partir da data da entrada em vigor desta Portaria, bem como sobre as regras para a representação das sequências de nucleotídeos e/ou de aminoácidos na Listagem de Sequências.

Art. 2º O requerente de pedido de patente que contenha em seu objeto uma ou mais sequências de nucleotídeos e/ou de aminoácidos, que sejam fundamentais para a descrição da invenção, deverá representá-las em uma Listagem de sequências, com vistas à aferição da suficiência descritiva, de que trata o art. 24 da Lei nº 9279 de 14 de maio de 1996 (doravante LPI).

Art. 3º A Listagem de sequências deverá ser apresentada ao INPI, como instrumento complementar ao relatório descritivo, através do sistema do Peticionamento Eletrônico, em arquivo eletrônico da Listagem de sequências gerado em formato texto TXT ou XML. A habilitação do campo de anexação da listagem durante o preenchimento do formulário de depósito substitui a apresentação adicional de Declaração de Veracidade.

Art. 4º A representação das sequências de nucleotídeos e/ou de aminoácidos na Listagem de sequências deverá seguir o Padrão OMPI ST.25, definido pela Organização Mundial da Propriedade Intelectual - OMPI, de acordo com as regras constantes do Anexo a esta Resolução.

§ 1º Devem ser incluídas na Listagem de sequências todas as sequências lineares de 4 (quatro) ou mais L-aminoácidos contínuos de um peptídeo ou de uma proteína e todas as sequências lineares que tenham 10 (dez) ou mais nucleotídeos contínuos, mesmo as que não tenham sido reivindicadas, como, por exemplo, sondas de PCR, desde que preencham as condições definidas neste parágrafo.

§ 2º As sequências ramificadas, as sequências com menos de 10 (dez) nucleotídeos, as sequências com menos de 4 (quatro) L-aminoácidos e as sequências de aminoácidos que contenham pelo menos um D-aminoácido, bem como as sequências compreendendo nucleotídeos ou aminoácidos diferentes dos que estão listados nas Tabelas 1, 2, 3 e 4, constantes do Anexo desta Resolução, devem ser incluídas no relatório descritivo do pedido de patente, não podendo constar da Listagem de sequências.

Art. 5º Da data de entrada em vigor desta Portaria até o dia 31/12/2021, a apresentação das sequências de nucleotídeos e/ou de aminoácidos na Listagem de sequências poderá, alternativa e voluntariamente, seguir o Padrão OMPI ST.26, definido pela OMPI, por meio de um único arquivo com formato XML, de acordo com as regras constantes no sítio do INPI na internet.

Parágrafo único. A Listagem de sequências em formato XML poderá ser criada, editada e verificada com a ferramenta WIPO Sequence, desenvolvida pela OMPI, que está disponível para download no sítio do INPI na internet, juntamente com o respectivo manual do usuário.

Art. 6º Após o carregamento do arquivo da listagem de Sequências, o sistema gerará o código de controle referente à listagem automaticamente e uma cópia em PDF do código será anexada na cópia do formulário recebida pelo depositante.

Art. 7º A apresentação do arquivo eletrônico da Listagem de sequências ao INPI deverá ser no ato do depósito do pedido de patente.

§ 1º Quando arquivo eletrônico da Listagem de sequências não for apresentado ao INPI no ato do depósito, poderá ser apresentado pelo requerente, independentemente de notificação ou exigência por parte do INPI, até a data do requerimento do exame do pedido de patente, de que trata o art. 33 da LPI, por meio de petição isenta do pagamento de retribuição.

§ 2º Quando a Listagem de sequências no formato de arquivo eletrônico não for apresentada nos prazos previstos no caput e no parágrafo primeiro deste artigo, o INPI formulará as exigências necessárias à regularização do pedido de patente, com vistas ao cumprimento do disposto Portaria, que deverão ser atendidas, nos termos e prazos da LPI.

Art. 8º Se a Listagem de sequências for corrigida subsequentemente a sua apresentação, de ofício ou a requerimento do requerente, este deverá apresentar ao INPI novo arquivo eletrônico da Listagem de sequências corrigida, observando as disposições desta Portaria, acompanhada do comprovante do recolhimento da retribuição correspondente ao ato processual.

Art. 9 As disposições desta Portaria aplicam-se ao pedido de patente oriundo de pedido internacional de patente depositado nos termos do Tratado de Cooperação em Matéria de Patentes - PCT, quando da sua entrada na fase nacional, apresentado ao INPI em conformidade com a legislação vigente.

Art. 10 A Listagem de sequências poderá ser adicionalmente apresentada em formato PDF, como parte integrante do pedido de patente.

§ 1º A Listagem de sequências que for adicionalmente apresentada no formato PDF quando do depósito do pedido de patente, deverá ser incluída após o relatório descritivo, sendo iniciada em uma página separada, sob o título Listagem de sequências.

§ 2º As páginas da Listagem de sequências de que trata o caput deverão ser numeradas de forma sequencial e independente, com algarismos arábicos, no centro da parte superior, entre 1 e 2 cm do limite da página.

Art. 11 Constará da Carta-Patente, além das informações e dos documentos de que trata o Art. 39 da LPI, o Código de Controle Alfanumérico referente à Listagem de sequências.

Parágrafo único. A Listagem de sequências referida no caput poderá ser acessada no site do INPI.

Art. 12 - Revoga-se a Resolução INPI/PR Nº 187 / 2017 de 27/04/2017.

Art. 13 - A presente Portaria entrará em vigor no dia 04 de janeiro de 2021.

Rio de Janeiro, 22 de dezembro de 2020.

CLÁUDIO VILAR FURTADO
Presidente

LIANE ELIZABETH CALDEIRA LAGE
Diretora de Patentes, Programas de Computador e
Topografias de Circuitos Integrados

Documento assinado eletronicamente por **CLAUDIO VILAR FURTADO**, **Presidente**, em 23/12/2020, às 10:49, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do Decreto nº 8.539, de 8 de outubro de 2015.

Documento assinado eletronicamente por **LIANE ELIZABETH CALDEIRA LAGE**, **Diretor(a)**, em 23/12/2020, às 12:38, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do Decreto nº 8.539, de 8 de outubro de 2015.

A autenticidade deste documento pode ser conferida no site http://sei.inpi.gov.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **0361044** e o código CRC **607CFD06**.

ANEXO I

REGRAS PARA APRESENTAÇÃO E REPRESENTAÇÃO DE SEQUÊNCIAS DE AMINOÁCIDOS E DE NUCLEOTIDEOS NA "LISTAGEM DE SEQUÊNCIAS" NO FORMATO OMPI ST.25

1 DAS DEFINIÇÕES**:**

1.1 Identificador de sequências é um número inteiro único que corresponde a SEQ ID NO: assinalada para cada sequência da listagem de sequências, sendo que a primeira sequência definida na "Listagem de Sequências", SEQ ID NO: 1, deve ser a sequência mais importante da invenção.

1.2 Identificador numérico é um número de três dígitos que representa um elemento específico de dados, alocado entre os símbolos < >.

1.3 Vocabulário linguisticamente neutro corresponde a um vocabulário padrão que se utiliza na listagem de sequências para representar os termos científicos no formato prescrito por provedores de dados de sequências (incluindo o nome científico, os qualificadores e seus valores em relação ao vocabulário, os símbolos das Tabelas 1, 2, 3 e 4 e as chaves de caracterização que figuram nas Tabelas 5 e 6).

1.4 Texto livre é a descrição textual das características de uma sequência em virtude do identificador numérico <223> (Outra informação), na qual se emprega um vocabulário distinto do vocabulário linguisticamente neutro definido no item 1.3.

2 DA REPRESENTAÇÃO DAS SEQUÊNCIAS BIOLÓGICAS NO FORMATO OMPI ST.25**:**

2.1. Cada sequência deverá ser assinalada com um identificador de sequência distinto. Os identificadores de sequências deverão ser iniciados com o número 1 e irão aumentando sequencialmente por números inteiros tais como "SEQ ID NO:1", "SEQ ID NO:2", "SEQ ID NO:3, etc.

2.2. No relatório descritivo, nas reivindicações e nos desenhos do pedido, as sequências representadas na listagem de sequências deverão ser referidas mediante o identificador de sequência precedido de "SEQ ID NO:".

2.3. As sequências de nucleotídeos e de aminoácidos deverão estar representadas por pelo menos uma das três possibilidades seguintes:

(i) uma sequência de nucleotídeos pura;

(ii) uma sequência de aminoácidos pura;

(iii) uma sequência de nucleotídeos e a correspondente sequência de aminoácidos.

2.4. Nas sequências representadas no formato especificado na opção (iii), a sequência de aminoácidos deverá ser adicionalmente revelada na listagem de sequências como uma sequência de aminoácidos pura e com um identificador de sequência diferente, composto por um número inteiro.

3. DO FORMATO E DOS SÍMBOLOS QUE DEVEM SER UTILIZADOS EM SEQUÊNCIAS DE NUCLEOTÍDEOS**:**

3.1 Toda sequência de nucleotídeos deverá ser representada unicamente por fita simples, no sentido 5' para 3' e da esquerda para a direita.

3.2 Toda sequência de nucleotídeos deverá ser representada por um máximo de 60 bases por linha, tendo um espaço entre cada grupo de 10 bases.

3.3 As bases das regiões codificadoras de uma sequência de nucleotídeos deverão figurar como tripletes (códons).

3.4 As bases de uma sequência de nucleotídeos deverão ser representadas usando o código de uma letra para os caracteres de nucleotídeos de sequência. Somente deverão ser usadas letras minúsculas, em conformidade com a listagem fornecida na Tabela 1.

3.5 As bases modificadas deverão ser representadas mediante as bases correspondentes não modificadas ou mediante o caractere "n" na própria sequência, caso a base modificada é uma das que figurem na Tabela 2.

4. DO FORMATO E DOS SÍMBOLOS QUE DEVEM SER UTILIZADOS EM SEQUÊNCIAS DE AMINOÁCIDOS**:**

4.1 Toda sequência de proteína ou de peptídeo deverá ser representada com um máximo de 16 aminoácidos por linha, deixando um espaço entre cada aminoácido.

4.2 As bases modificadas deverão ser representadas mediante as bases correspondentes não modificadas ou mediante o caractere "n" na própria sequência, caso a base modificada é uma das que figurem na Tabela 2.

4.3 A numeração dos aminoácidos deverá ser iniciada no primeiro aminoácido da sequência com o número 1.

4.4 Alternativamente, os aminoácidos que precedem a proteína madura, por exemplo, as pré-sequências, as pró-sequências e as pré-pró-sequências, assim como as sequências sinal, quando existentes, poderão ter números negativos, contados em forma regressiva, a partir do aminoácido adjacente ao número 1.

4.5 Não se empregará o zero (0) quando a numeração dos aminoácidos empregar números negativos para distinguir a proteína madura.

4.6 Toda sequência de aminoácidos composta por um ou mais segmentos não contínuos de uma sequência maior ou de segmentos de sequências diferentes, deverá ser numerada como uma sequência distinta e com um identificador de sequência diferente.

4.7 Os aminoácidos de uma sequência de proteína ou de peptídeo deverão ser representados no sentido do grupamento amino para o grupamento carboxila e da esquerda para a direita.

4.8 Os aminoácidos deverão ser representados utilizando o código de três letras, sendo a primeira letra uma letra maiúscula, em conformidade com a listagem dada na Tabela 3.

5 DOS ELEMENTOS DE DADOS OBRIGATÓRIOS**:**

5.1. A listagem de sequências deverá incluir, em adição a, e imediatamente antes da sequência de nucleotídeos e/ou aminoácidos, os seguintes elementos de informação (elementos de dados obrigatórios):

<110>	Nome do requerente
<120>	Título da invenção
<160>	Número total de SEQ ID NOs
<210>	SEQ ID NO: #
<211>	Comprimento
<212>	Tipo
<213>	Organismo
<400>	Sequência

Quando o nome do requerente (identificador numérico <110>) estiver escrito em caracteres outros que não os pertencentes ao alfabeto latino, também deverá aparecer em caracteres do alfabeto latino, seja como uma simples transliteração do nome ou através da sua tradução para o inglês.

5.2. Se for empregado na sequência o caractere “n” ou Xaa”, ou uma base modificada, ou um L-aminoácido modificado ou pouco comum, os seguintes elementos de dados serão obrigatórios:

<220>>	Característica
<221>	Nome/chave
<222>	Localização
<223>	Outra informação

5.3. Se o organismo (identificador numérico <213>) é uma “Sequência artificial” ou “Desconhecida”, os seguintes elementos de dados são obrigatórios:

<220>	Característica
<223>	Outra informação

5.4. Quando uma listagem de sequências é apresentada em conjunto com o pedido de patente no ato de seu depósito ou em qualquer momento antes da designação de um número de depósito ao mesmo, o seguinte elemento de dados deverá estar incluído obrigatoriamente na listagem de sequências:

<130>	Número de referência pessoal (indicado pelo requerente)

5.5. Quando uma listagem de sequências é apresentada em resposta a uma exigência emitida por este INPI ou a qualquer momento após a designação de um número de depósito, os seguintes elementos de dados deverão estar obrigatoriamente incluídos na “Listagem de Sequências”:

<140>	Número do pedido de patente em trâmite
<141>	Data de depósito do pedido de patente

5.6. Além dos elementos de dados identificados acima, quando uma listagem de sequências é apresentada em relação a um pedido na qual se reivindica a prioridade de um pedido anterior, os seguintes elementos de dados deverão constar na “Listagem de Sequências”:

<150>	Pedido de patente anterior (documento de prioridade)

<151>	Data de depósito do pedido de patente anterior (dia/mês/ano)

6. DA APRESENTAÇÃO DAS CARACTERÍSTICAS**:**

6.1. Quando características da sequência são apresentadas (ou seja, identificador numérico <220>), as mesmas deverão ser descritas mediante as "chaves de caracterização" definidas nas Tabelas 5 e 6.

7. TEXTO LIVRE**:**

7.1 A utilização do texto livre deverá estar limitada a uns poucos termos curtos que sejam indispensáveis para o entendimento da sequência.

7.2. Cada elemento de dados não excederá a quatro linhas com um máximo de 65 caracteres por linha.

7.3. Qualquer informação adicional deverá ser incluída na parte principal do relatório descritivo.

IDENTIFICADORES NUMÉRICOS OBRIGATÓRIOS

identificador numérico	**Descrição do identificador numérico**	**Comentário**
<110>	Nome do requerente	Quando o nome do requerente estiver escrito em caracteres diferentes dos que compõem o alfabeto latino, também deverá ser indicado em caracteres do alfabeto latino, seja como simples transliteração ou mediante a sua tradução para o inglês; havendo mais de um requerente, listar um nome por linha.
<120>	Título da invenção	Em língua vernácula
<130>	Número de referência do pedido	Obrigatório somente nas condições especificadas pelo item 5.4.
<140>	Pedido de patente em trâmite	Obrigatório somente nas condições especificadas pelo item 5.5.
<141>	Data de depósito do pedido de patente em trâmite	Obrigatório somente nas condições especificadas pelo item 5.5.
<150>	Pedido de patente anterior (prioridade)	Obrigatório somente na condição especificada pelo item 5.6.
<151>	Data de depósito do pedido de patente anterior (prioridade)	Obrigatório somente na condição especificada pelo item 5.6.
<160>	Número de SEQ ID NOs	Inclui o número total de SEQ ID NOs compreendidas na listagem de sequências
<210>	Informação sobre a SEQ ID NO: #	A resposta deverá estar composta por um número inteiro que represente a SEQ ID NO mostrada
<211>	Comprimento	Comprimento da sequência expressa em número de pares de bases ou de resíduos de aminoácidos
<212>	Tipo	Tipo de molécula DNA/RNA/PROTEÍNA que é mostrada na SEQ ID NO: #, ou seja, DNA, RNA ou PRT (proteína); se a sequência de nucleotídeos contiver fragmentos de DNA e de RNA, o tipo será "DNA"; além disso, a molécula combinada de DNA/RNA também deverá ser objeto de descrição na seção de características <220> a <223>.

<213>	Organismo	Gênero e espécie (ou seja, o nome científico) ou “Sequência Artificial” (*Artificial Sequence*) ou “Desconhecido” (*Unknown*); adicionalmente, a sequência artificial ou o organismo desconhecido deverá ser também objeto de descrição na seção de características <220> a <223>
<220>	Característica	Obrigatório somente nas condições especificadas pelos itens 5.2 e 5.3. Caso contrário, deixe em branco.
<221>	Nome/chave	Obrigatório somente na condição especificada pelo item 5.2.
<222>	Localização	Obrigatório somente na condição especificada pelo item 5.2.
<223>	Outras informações	Obrigatório somente nas condições especificadas pelos itens 5.2 e 5.3.
<400>	Sequência	O elemento SEQ ID NO: deve ir depois do identificador numérico e deve figurar na linha anterior a sequência de fato

TABELA 1: LISTAGEM DE NUCLEOTÍDEOS

Símbolo	**Significado**	Origem da designação
a	a	adenina
g	g	guanina
c	c	citosina
t	t	timina
u	u	uracila
r	g ou a	purina
y	t/u ou c	pirimidina (*pyrimidine*)
m	a ou c	amino
k	g ou t/u	ceto (*keto*)
s	g ou c	interações fortes (*strong interactions)* 3 (três) pontes de hidrogênio
w	a ou t/u	interações fracas (*weak interactions*) 2 (duas) pontes de hidrogênio
b	g ou c ou t/u	que não seja a
d	a ou g ou t/u	que não seja c

h	a ou c ou t/u	que não seja g
v	a ou g ou c	que não seja t e nem u
n	a ou g ou c ou t/u, desconhecido ou outro	qualquer (*any*)

TABELA 2: LISTAGEM DE NUCLEOTÍDEOS MODIFICADOS

Símbolo	**Significado**
ac4c	4-acetilcitidina
chm5u	5-(carboxihidroximetil)uridina
cm	2’-O-metilcitidina
cmnm5s2u	5-carboximetilaminometil-2-tiouridina
cmnm5u	5-carboximetilaminometiluridina
d	dihidrouridina
fm	2’-O-metilpseudouridina
gal q	beta, D-galactosilqueosine
gm	2’-O-metilguanosina
i	Inosina
i6a	N6-isopenteniladenosina
m1a	1-metiladenosina
m1f	1-metilpseudouridina
m1g	1-metilguanosina
m1i	1-metilinosina
m22g	2,2-dimetilguanosina
m2a	2-metiladenosina
m2g	2-metilguanosina
m3c	3-metilcitidina
m5c	5-metilcitidina
m6a	N6-metiladenosina
m7g	7-metilguanosina
mam5u	5-metilaminometiluridina

mam5s2u	5-metoxiaminometil-2-tiouridina
man q	beta, D-manosilqueosina
mcm5s2u	5-metoxicarbonilmetil-2-tiouridina
mcm5u	5-metoxicarbonilmetiluridina
mo5u	5-metoxiuridina
ms2i6a	2-metiltio-N6-isopenteniladenosina
ms2t6a	N-((9-beta-D-ribofuranosil-2-metiltiopurina-6-il)carbamoil)treonina
mt6a	N-((9-beta-D-ribofuranosilpurina-6-il)N-metilcarbamoil)treonina
mv	5-metoxicarbonilmetoxiuridina
o5u	uridina-5-ácido oxiacético
osyw	wybutoxosina
p	pseudouridina
q	queosina
s2c	2-tiocitidina
s2t	5-metil-2-tiouridina
s2u	2-tiouridina
s4u	4-tiouridina
t	5-metiluridina
t6a	N-((9-beta-D-ribofuranosilpurina-6-il)-carbamoil)treonina
tm	2’-O-metil-5-metiluridina
um	2’-O-metiluridina
yw	wybutosina
x	3-(3-amino-3-carboxi-propil)uridina, (acp3)u

TABELA 3: LISTAGEM DE AMINOÁCIDOS

Símbolo	**Significado**
Ala	Alanina
Cys	Cisteína
Asp	Ácido Aspártico
Glu	Ácido Glutâmico

Phe	Fenilalanina
Gly	Glicina
His	Histidina
Ile	Isoleucina
Lys	Lisina
Leu	Leucina
Met	Metionina
Asn	Asparagina
Pro	Prolina
Gln	Glutamina
Arg	Arginina
Ser	Serina
Thr	Treonina
Val	Valina
Trp	Triptofano
Tyr	Tirosina
Asx	Asp ou Asn
Glx	Glu ou Gln
Xaa	desconhecido ou outro

TABELA 4: LISTAGEM DE AMINOÁCIDOS MODIFICADOS OU POUCO USUAIS

Símbolo	**Significado**
Aad	Ácido 2-aminoadípico
bAad	Ácido 3-aminoadípico
bAla	beta-Alanina, ácido beta-aminopropiônico
Abu	Ácido 2-aminobutírico
4Abu	Ácido 4-aminobutírico, ácido piperidínico
Acp	Ácido 6-aminocapróico
Ahe	Ácido 2-aminoheptanóico
Aib	Ácido 2-aminoisobutírico

bAib	Ácido 3-aminoisobutírico
Apm	Ácido 2-aminopimélico
Dbu	Ácido 2,4 diaminobutírico
Des	Desmosina
Dpm	Ácido 2,2'-diaminopimélico
Dpr	Ácido 2,3-diaminopropiônico
EtGly	N-etilglicina
EtAsn	N-etilasparagina
Hyl	Hidroxilisina
aHyl	alo-Hidroxilisina
3Hyp	3-Hidroxiprolina
4Hyp	4-Hidroxiprolina
Ide	Isodesmosina
alle	alo-Isoleucina
MeGly	N-metilglicina, sarcosina
Melle	N-metilisoleucina
MeLys	6-N-metillisina
MeVal	N-metilvalina
Nva	Norvalina
Nle	Norleucina
Orn	Ornitina

TABELA 5: LISTAGEM DAS CHAVES DE CARACTERIZAÇÃO DE SEQUÊNCIAS DE NUCLEOTÍDEOS

Chave	**Descrição**
allele (alelo)	Existência de indivíduos ou estirpes relacionadas que contém formas estáveis e diferentes do mesmo gene e que diferem da sequência apresentada nesta localização (e talvez em outras)
attenuator (atenuador)	1) região do DNA onde ocorre controle da terminação da transcrição que controla a expressão de certos operadores bacterianos; 2) segmento de sequência localizado entre o promotor e o primeiro gene estrutural que causa terminação parcial da transcrição

*C*_region (região-C)	Região constante das cadeias leve e pesada das imunoglobulinas e das cadeias alfa, beta e gama do receptor de linfócitos T; inclui um ou mais exons, dependendo da cadeia em particular
CAAT_signal (sinal CAAT)	Região CAAT *box*; parte de uma sequência conservada situada à cerca de 75 pares de bases a montante do local de iniciação das unidades de transcrição eucarióticas e que pode estar envolvida na ligação da RNA polimerase sequência consenso= GG (C ou T) CAATCT
CDS (sequência codificadora)	Sequência codificadora (*coding sequence*); sequência de nucleotídeos que se corresponde com a sequência de aminoácidos de uma proteína (a localização inclui o códon de terminação); contém a tradução conceptual dos aminoácidos
conflict (conflito)	Determinações independentes da "mesma" sequência diferem neste local ou nesta região
D-loop (alça de deslocamento)	Alça de deslocamento (*Displacement loop*); região do DNA mitocondrial na qual uma sequência curta de RNA fita simples é pareada com uma das fitas do DNA, deslocando nesta região a outra fita de DNA pareada; também usada para descrever o deslocamento de uma região de fita simples em um DNA duplex por um invasor fita simples, na reação catalisada pela proteína RecA
D-segment (segmento de diversidade)	Segmento de diversidade (*Diversity segment*) da cadeia pesada das imunoglobulinas e da cadeia pesada do receptor de linfócitos T
enhancer (acentuador)	*Enhancer* ou acentuador é uma sequência que aumenta a utilização de (certos) promotores eucarióticos situados na mesma fita de DNA (efeito em cis) e cuja ação pode efetuar-se com independência da orientação e da localização (5' ou 3') em relação ao promotor
exon (éxon)	Região do genoma que codifica para a porção do RNA mensageiro processado (*spliced mRNA*); pode conter a região 5'UTR, todas as sequências codificadoras (CDS) e a região 3'UTR
GC_signal (sinal GC)	Região GC *box*; região conservada rica em GC e localizada antes do ponto de iniciação das unidades de transcrição eucarióticas e que pode adotar a forma de múltiplas cópias e produzir-se em ambos os sentidos (5' ou 3') sequência consenso= GGGCGG
gene (gene)	Região de interesse biológico identificada como sendo um gene e para a qual foi designado um nome; ácido nucléico codificador
iDNA (DNA de intervenção)	DNA de intervenção (*intervening DNA*); DNA que é eliminado em diferentes tipos de recombinação
intron	

(íntron)	Segmento de DNA que é transcrito, porém logo removido da nova molécula de RNA pelo processo de *splicing* do RNA, ocasionando junção dos éxons que flanqueiam os íntrons
J_segment (segmento de ligação)	Segmento de ligação (*Joining segment*) das cadeias leve e pesada das imunoglobulinas e das cadeias alfa, beta e gama do receptor de célula T
LTR *(*sequências repetitivas longas)	LTRs (*Long Terminal Repeat*) são sequências repetitivas longas encontradas em cada extremidade (5' e 3') de uma sequência tal como a que é tipicamente encontrada nos retrovírus
mat_peptide (sequência codificadora de um peptídeo)	Sequência codificadora de um peptídeo ou de uma proteína madura; sequência codificadora do peptídeo ou da proteína em sua condição madura ou final, seguida de modificação pós-tradução; a localização não inclui o códon de terminação (diferentemente da CDS correspondente)
misc_binding	Região em um ácido nucléico que se liga covalentemente ou não com outra molécula e que não pode ser descrito por qualquer outra chave de ligação (primer_bind ou protein_bind)
misc_difference	A sequência caracterizada é diferente nesta posição, daquela apresentada na entrada e não pode ser descrita por nenhuma outra chave de diferença (conflict, unsure, old_sequence, mutation, variation, allele ou modified_base)
misc_feature	Região de interesse biológico que não pode ser descrita por nenhuma outra chave de característica; uma característica nova ou pouco comum
misc_recomb	Sítio de qualquer recombinação generalizada, sitio-especifica ou replicativa, por onde se produz a excisão e ligação de DNA duplex e que não pode ser descrita por nenhuma outra chave de recombinação (iDNA ou virion) e nem por qualificadores da chave de origem (/insertion_seq, /transposon, /proviral)
misc_RNA	Qualquer porção transcrita ou RNA que não pode ser definida por nenhuma outra chave de RNA (prim_transcript, precursor_RNA, mRNA, 5'clip, 3'clip, 5'UTR, exon, CDS, sig_peptide, transit_peptide, mat_peptide, intron, polyA_site, rRNA, tRNA, scRNA ou snRNA)
misc_signal	Qualquer região que contenha um sinal que controla ou modifica uma função ou expressão de um gene, que não pode ser descrito por nenhuma outra chave de sinal (promoter, CAAT_signal, TATA_signal, -35_signal, -10_signal, GC_signal, RBS, polyA_signal, enhancer, attenuator, terminator ou rep_origin)
misc_structure	Qualquer conformação ou estrutura secundária ou terciária que não pode ser descrita por nenhuma outra chave de estrutura (stem_loop ou D-loop)
modified_base (nucleotídeo modificado)	O nucleotídeo indicado é um nucleotídeo modificado e deve ser substituído pela molécula indicada (que figura no valor qualificador de mod_base)

mRNA (RNA mensageiro)	RNA mensageiro; inclui a região 5' não traduzida (5'UTR), a sequência codificadora (CDS, exon) e a região 3' não traduzida (3'UTR)
mutation (mutação)	Uma estirpe relacionada apresenta uma alteração brusca e não transmissível na sequência, nesta localização
região N (N_region)	Região de inserção de nucleotídeos adicionais entre os segmentos reordenados das imunoglobulinas
old_sequence (prévia sequência)	A sequência apresentada é uma versão revisada de uma prévia sequência nesta localização
polyA_signal (sinal de poliadenilação)	Região indispensável de reconhecimento para clivagem por uma endonuclease seguida por poliadenilação de uma porção transcrita de RNA sequência consenso= AATAAA
polyA_site (sítio de poliadenilação)	Região de um transcrito de RNA no qual se adicionam resíduos de adenina por poliadenilação pós-transcricional
precursor_RNA (RNA precursor)	Precursor de RNA, qualquer RNA imaturo; pode incluir a região cortada em 5' (5'clip), a região 5' não traduzida (5'UTR), as sequências codificadoras (CDS, exon), as sequências intervenientes (intron), a região 3' não traduzida (3'UTR) e a região cortada em 3' (3'clip)
prim_transcript (transcrito primário)	Transcrito primário (inicial, não processado); inclui a região cortada em 5' (5'clip), a região 5' não traduzida (5'UTR), as sequências codificadoras (CDS, exon), as sequências intervenientes (intron), a região 3' não traduzida (3'UTR) e a região cortada em 3' (3'clip)
primer_bind (região de ligação de um iniciador)	Região de ligação não covalente de um iniciador (*primer*) na iniciação da replicação, da transcrição ou da transcrição reversa; inclui as regiões para iniciadores sintéticos, por exemplo, os que são usados na reação em cadeia da polimerase (PCR)
promoter (promotor)	Região de uma molécula de DNA na qual se liga a RNA polimerase para iniciar a transcrição
protein_bind (ligação de proteína)	Região de ligação não covalente de proteínas em um ácido nucléico
RBS	

(sítio de ligação de ribossomo)	Região de ligação do ribossomo (*ribosome binding site*)
repeat_region (região repetitiva)	Região do genoma que contém unidades de repetição
repeat_unit (unidade de repetição)	Elemento (unidade de repetição) que se repete na *repeat_region*
rep_origin (origem de replicação)	Origem de replicação; região onde se inicia a duplicação de um ácido nucléico para obter duas cópias idênticas
rRNA (rRNA)	RNA ribossomal maduro; RNA que compõe a partícula ribonucleoprotéica (ribossomo) que sintetiza proteínas a partir de aminoácidos
S_region (região S)	Região de mudança (*switch region*) das cadeias pesadas das imunoglobulinas; envolvida no rearranjo do DNA que codifica para a cadeia pesada levando à expressão de uma classe diferente de imunoglobulina por um mesmo linfócito B
satellite (satélite)	Múltiplas repetições em *tandem* (idênticas ou parecidas) de uma unidade de repetição básica curta; muitas delas têm uma composição de bases ou uma outra propriedade diferente do genoma em geral, o que permite separá-las do resto do DNA genômico (banda principal)
scRNA (RNA citoplasmático pequeno)	RNA citoplasmático de tamanho pequeno (*small cytoplasmic RNA*); uma das diversas pequenas moléculas de RNA presentes no citoplasma e (algumas vezes) no núcleo de uma célula eucariótica
sig_peptide (peptídeo sinal)	Sequência codificadora para um peptídeo sinal; sequência codificadora do domínio amino-terminal de uma proteína secretada; este domínio está envolvido na integração do polipeptídeo nascente na membrana; sequência leader
snRNA (RNA nuclear pequeno)	RNA nuclear de tamanho pequeno (*small nuclear RNA*); qualquer uma das muitas espécies de RNA pequeno que estão confinadas no núcleo; vários dos snRNA estão envolvidos em *splicing* ou em outras reações de processamento de RNA
source (fonte)	Identifica a fonte biológica do intervalo de sequência especificamente indicado; esta chave é obrigatória; cada entrada deve estar composta por no mínimo, de uma chave única de fonte englobando a sequência inteira; é permitido o uso de mais de uma chave de fonte por sequência
stem_loop (alça em forma de grampo)	Alça em forma de grampo (*hairpin*); região de dupla hélice formada pelo pareamento de bases entre sequências complementares adjacentes (invertidas) que pertencem a uma mesma fita de RNA ou de DNA (pareamento intramolecular)

STS (região marcadora de DNA)	Regiões marcadoras na sequência (*Sequence Tagged Site*); trata-se de sequências curtas de DNA que ocorrem uma única vez no genoma humano e cuja posição exata e ordem de bases, uma vez conhecidas, identificam um local no genoma, sendo detectadas por PCR; o mapa de uma região do genoma pode efetuar-se determinando a ordem de uma série de STS
TATA_signal (sinal TATA)	TATA-*box*; Goldberg-Hogness *box*; é um heptâmero conservado rico em A•T, situado a cerca de 25 pares de bases antes do sítio de iniciação de cada unidade transcrita pela RNA polimerase II das células eucarióticas; pode estar envolvido no posicionamento da enzima para a iniciação correta da transcrição sequência consensual= TATA(A ou T)A(A ou T)
terminator (terminador)	*Terminator* ou terminador; sequência de DNA localizada no final do transcrito ou adjacente a um promotor e que faz com que a RNA polimerase termine a transcrição; também pode ser o sítio de ligação da proteína repressora
transit_peptide (peptídeo de trânsito)	Sequência codificadora para um peptídeo de trânsito; sequência codificadora do domínio amino-terminal de uma proteína de organela codificada no núcleo; este elemento está envolvido na importação pós-tradução da proteína para dentro da organela
tRNA (RNA transportador)	RNA de transferência maduro, RNA de tamanho pequeno (75-85 bases) que media a tradução de uma sequência de ácido nucléico em uma sequência de aminoácidos
unsure (incerto)	O autor não está seguro sobre a exatidão da sequência nesta região
V_region (região V)	Região variável das cadeias leve e pesada das imunoglobulinas e das cadeias alfa, beta e gama do receptor de linfócitos T; codifica para a região variável na extremidade amino-terminal; pode estar composta por: V_segment, D_segment, N_region e J_segment
V_segment (segmento V)	Segmento variável das cadeias leve e pesada das imunoglobulinas e das cadeias alfa, beta e gama do receptor de linfócitos T; codifica para a maior parte da região variável (V_region) e para os últimos aminoácidos do peptídeo lider (*leader peptide*)
variation (variante)	Existência de uma estirpe relacionada que contém mutações estáveis do mesmo gene (por exemplo, RFLP, polimorfismos, etc) e que diferem da sequência apresentada nesta localização (e talvez em outras)
3'clip	Região na extremidade 3' de um RNA precursor que é cortado durante processamento
3'UTR	Região na extremidade 3' (posterior ao códon de terminação) de um RNA maduro que não se traduz em proteína

5'clip	Região na extremidade 5' de um RNA precursor que é cortado no processamento
5'UTR	Região na extremidade 5' (anterior ao códon de terminação) de um RNA maduro que não se traduz em proteína
-10_signal (sinal -10)	Sequência -10 (*pribnow box*); sequência conservada centrada aproximadamente 10 pares de bases antes do sítio de início da transcrição de um gene bacteriano e que pode participar na ligação da RNA polimerase sequência consenso= TAtAaT
-35_signal (sinal –35)	Sequência -35; sequência centrada aproximadamente 35 pares de bases antes do sítio de início da transcrição de um gene bacteriano sequência consenso= TTGACa ou TGTTGACA

TABELA 6: LISTAGEM DE CHAVES DE CARACTERIZAÇÃO DE SEQUÊNCIAS DE AMINOÁCIDOS

Chave	**Descrição**
CONFLICT (CONFLITO)	Diferentes documentos reportam diferentes sequências
VARIANT (VARIANTE)	Os autores assinalam que existem variações da sequência
VARSPLIC (VARIANTE DE EDIÇÃO)	Descrição das variações da sequência produzidas por um *splicing* alternativo
MUTAGEN (SÍTIO ALTERADO POR MUTAÇÃO)	Sítio que foi experimentalmente alterado
MOD_RES (RESÍDUO PÓS-MODIFICADO)	Modificação pós-tradução de um resíduo
ACETYLATION (ACETILAÇÃO)	Acetilação na extremidade amino-terminal ou outra
AMIDATION (AMIDAÇÃO)	Amidação geralmente na extremidade carboxi-terminal de um peptídeo maduro e ativo

BLOCKED (SÍTIO BLOQUEADO)	Grupo de bloqueio indeterminado na extremidade amino-terminal ou carboxi-terminal
FORMYLATION (FORMILAÇÃO)	Formilação da metionina da extremidade amino-terminal
GAMMA-CARBOXYGLUTAMIC ACID HYDROXYLATION (HIDROXILAÇÃO ÁCIDO GAMA-CARBOXIGLUTÂMICO)	da asparagina, do ácido aspártico, da prolina ou da lisina
METHYLATION (METILAÇÃO)	Metilação geralmente da lisina ou da arginina
PHOSPHORYLATION (FOSFORILAÇÃO)	Fosforilação da serina, da treonina, da tirosina, do ácido aspártico ou da histidina
PYRROLIDONE CARBOXYLIC ACID (ÁCIDO CARBOXI PIRROLIDÔNICO)	Glutamato amino-terminal que formou uma lactama cíclica interna
SULFATATION (SULFATAÇÃO)	Sulfatação geralmente da tirosina
LIPID (LIPÍDIO)	Ligação covalente de um fragmento lipídico
MYRISTATE (MIRISTATO)	Grupo miristato unido por uma ligação amida a um resíduo de glicina da extremidade amino-terminal da forma madura de uma proteína ou de um resíduo interno de lisina
PALMITATE (PALMITATO)	Grupo palmitato unido por uma ligação tioéter a um resíduo de cisteína ou por uma ligação éster a um resíduo de serina ou de treonina
FARNESYL (FARNESIL)	Grupo farnesil ligado por uma ligação tioéter a um resíduo de cisteína
GERANYL-GERANYL (GERANIL-GERANIL)	Grupo geranil-geranil ligado por uma ligação tioéter a um resíduo cisteína
GPI-ANCHOR (GRUPO GLICOSIL-FOSFATIDILINOSITOL ANCORADO)	Grupo glicosil-fosfatidilinositol (GPI) unido a um grupo alfa-carboxila do resíduo carboxi-terminal da forma madura de uma proteína
N-ACYL DIGLYCERIDE	Cisteína amino-terminal da forma madura de uma lipoproteína de procarioto unida por uma ligação amida a um ácido graxo e um grupo

(*N-ACIL DICLIiCERÍDEO*)	gliceril, na qual dois ácidos graxos estão unidos por ligação éster
DISULFID (PONTE DISSULFETO)	Ponte dissulfeto; os extremos "DE" ("*FROM*") e "PARA" ("*TO*") representam os dois resíduos que estão ligados por uma ponte dissulfeto intra-cadeia peptídica; se os extremos "DE ("*FROM*") e "PARA" ("*TO*") são idênticos, a ponte dissulfeto é uma ligação inter-cadeia peptídica e o campo descritivo indica a natureza das ligações cruzadas (*cross-link*)
THIOLEST (LIGAÇÃO TIOÉSTER)	Ligação tioéster; os extremos "DE ("*FROM*") e "PARA" ("*TO*") representam os dois resíduos que estão unidos pela ligação tioéster
THIOETH (LIGAÇÃO TIOÉTER)	Ligação tioéter; os extremos "DE ("*FROM*") e "PARA" ("*TO*") representam os dois resíduos que estão unidos pela ligação tioéter
CARBOHYD (SÍTIO DE GLICOSILAÇÃO)	Sítio de glicosilação; a natureza do carboidrato (se conhecido) está indicada no campo descritivo
METAL (SÍTIO DE LIGAÇÃO DE METAL)	Sítio de ligação para um íon de metal; no campo descritivo é indicada a natureza do metal
BINDING (SÍTIO DE LIGAÇÃO)	Sítio de ligação para qualquer grupo químico (coenzima, grupo prostético, etc.); no campo descritivo é indicada a natureza química do grupo
SIGNAL (SINAL)	Extensão de uma sequência-sinal (pré-peptídeo)
TRANSIT (TRÂNSITO)	Extensão de um peptídeo de trânsito (mitocondrial, cloroplástico ou destinado para microssoma)
PROPEP (PROPEP)	Extensão de um pró-peptídeo
CHAIN (CADEIA)	Extensão da cadeia polipeptídica na proteína madura
PEPTIDE (PEPTÍDEO)	Extensão de um peptídeo ativo liberado
DOMAIN (DOMÍNIO)	Extensão de um domínio de interesse na sequência; no campo descritivo é indicada a natureza deste domínio
CA_BIND (SÍTIO DE LIGAÇÃO DE CÁLCIO)	Extensão de uma região de ligação de cálcio

TRANSMEM (TRANSMEMBRANA)	Extensão de uma região transmembrana
ZN_FING (MOTIVO DEDO DE ZINCO)	Extensão de uma região contendo o motivo dedo de zinco (*zinc finger)*
SIMILAR (SIMILAR)	Extensão da similaridade de uma região com uma outra sequência protéica; no campo descritivo são indicadas informações detalhadas sobre esta sequência
REPEAT (SEQUÊNCIA INTERNA REPETITIVA)	Extensão de uma sequência interna repetitiva
HELIX (HÉLICE)	Estrutura secundária: Hélices, por exemplo, a alfa-hélice, a hélice 310 ou a hélice Pi
STRAND (FITA)	Estrutura secundária: folha beta (folha-b), por exemplo, folha beta-pregueada unida por pontes de hidrogênio, o resíduo isolado em uma ponte beta
TURN (VOLTA)	Estrutura secundária: voltas (*turns*), por exemplo, voltas mantidas por pontes de hidrogênio (voltas de 3, 4 ou 5 resíduos de aminoácidos)
ACT_SITE (SÍTIO ATIVO)	Aminoácidos envolvidos na atividade de uma enzima
SITE (SÍTIO)	Qualquer outro sítio de interesse na sequência
INIT_MET (INICIA COM METIONINA)	A sequência começa com uma metionina de iniciação
NON_TER (NÃO TERMINAL)	O resíduo em uma extremidade da sequência não é o resíduo terminal; se aplicado à posição 1, significa que a primeira posição não é a posição amino-terminal da molécula completa; se aplicado para a última posição, significa que esta posição não é a posição carboxi-terminal da molécula completa; não há nenhum campo descritivo para esta chave
NON_CONS (NÃO CONSECUTIVOS)	Resíduos não consecutivos; indica que dois resíduos de uma sequência não são consecutivos e que existem vários resíduos não sequenciados entre eles
UNSURE (INCERTO)	Zonas de incertezas na sequência; usado para descrever as regiões da sequência para as quais os autores não estão certos de sua definição

8. DOS ELEMENTOS DE DADOS NÃO OBRIGATÓRIOS**:**

8.1 Todos os elementos de dados citados a seguir são facultativos de comporem a “Listagem de Sequências”:

<170>	Programa de computador usado para gerar a listagem de sequências
<300>	Informações sobre publicação; havendo várias publicações, repita a seção para cada publicação relevante
<301>	Autores, especifique um nome por linha, preferencialmente no formato: sobrenome, outros nomes e/ou iniciais
<302>	Título da publicação
<303>	Nome do periódico no qual se publicaram os dados
<304>	Volume do periódico no qual se publicaram os dados
<305>	Número do periódico no qual se publicaram os dados
<306>	Número das páginas do periódico no qual se publicaram os dados
<307>	Data do periódico no qual se publicaram os dados; usar formato Dia/Mês/Ano
<308>	Número de acesso assinalado pela base de dados, incluindo o nome da base de dados
<309>	Data de entrada na base de dados (dia/mês/ano)
<310>	Número do documento de patente, unicamente para as patentes citadas
<311>	Data de submissão do documento de patente, unicamente para as patentes citadas (dia/mês/ano)
<312>	Data de publicação do documento de patente; unicamente para as patentes citadas (dia/mês/ano)
<313>	Resíduos relevantes na SEQ ID NO: #: DE (from)_PARA (to)