C

WIPO
世界
知识产权
组织

CWS/10/13
原文：英文
日期：2022 年 9 月 20 日

产权组织标准委员会（CWS）

第十届会议
2022 年 11 月 21 日至 25 日，日内瓦

关于修订产权组织标准 ST.26 的提案

国际局编拟的文件

概　述

1.　序列表工作队提出对产权组织标准 ST.26 的当前版本进行修订，以更正工作队指出的编辑性错误，并在附件六中加入三个新示例。在产权组织标准委员会（CWS）第九届会议上，国际局注意到主管局对进一步修订产权组织 ST.26 的关切，并就如何管理这些修改提供了指导。

背　景

2.　在 2016 年 3 月举行的第四届会议续会上， CWS 通过了产权组织标准 ST.26（见文件 CWS/4BIS/16 第 52 段），题为“用 XML（可扩展标记语言）表示核苷酸和氨基酸序列表的建议标准”。序列表工作队负责在第 44 号任务的框架下进行产权组织 ST.26 的修正，该任务说明如下：

“为国际局提供支持，提供用户对 ST.26 编著和验证软件工具的要求和反馈意见；在对《PCT 行政规程》进行相应修订的工作上，为国际局提供支持；并且为产权组织标准 ST.26 编制必要的修订。”

3.　CWS 第五、第六、第七和第八届会议上批准了对产权组织 ST.26 的修订，分别形成了第 1.1、1.2、1.3 和 1.4 版（见文件 CWS/5/22 第 41 段、文件 CWS/6/34 第 112 段、文件 CWS/7/29 第 133 段和文件 CWS/8/24 第 49 段）。最近，在 2021 年 11 月召开的第九届会议上，CWS 批准了若干编辑上的和实质性的修订，并作为产权组织 ST.26 的最新版本第 1.5 版予以公布（见文件 CWS/9/25 第 63 段）。这包括为与国际核酸序列数据库联盟（INSDC）的要求保持一致而做出的修改。

4. 在第九届会议期间，各代表团对所有成员国已实施产权组织 ST.26 后，修订产权组织 ST.26 可能产生的影响表示关切，例如其信息技术系统需要重新调整（见文件 CWS/9/25 第 62 段）。国际局表示将在第十届会议上就此提供指导。

5. 在注意到第九届会议作出修订后，所提供的序列编号有若干引用错误和不正确的特征位置后，对附件六和附件七的附录进行了若干编辑性修改。这些修改于 2022 年 5 月做出并公布。

产权组织标准 ST.26 的拟议修订

6. 序列表工作队编拟了产权组织 ST.26 的修订提案，供 CWS 审议并酌情批准。凡注明删除线的内容表示修改，加粗内容表示增加。

7. 对产权组织标准 ST.26 主体部分的拟议修改可总结如下：

(a) 鉴于一些知识产权局不允许在提交的电子文件的文件名中使用下划线‘_’，因此更新第 44 段以改进所提供的 filename 值，使其包含‘-’来代替‘_’：filename=“US11-405455-SEQL.xml”；

(b) 更新第 45 段，将 ApplicationNumberText 的定义改进为：‘The application number as provided by the office of filing (e.g., PCT/IB2013/099999).’；

(c) 更新第 73 段，在第一句的后半部分加入以下短语，以确保与第六部分的说明相一致：‘Section 8 provides the exclusive listing of qualifiers, **and their specified value formats, if any,** for each amino acid sequence feature key.’；

(d) 更新第 76 段，为‘id’属性提供一个新定义，用于语种相关的自由文本限定符，作为所含表格中新的第四行的一部分。该说明如下：‘A qualifier with language dependent free text may be uniquely identified by using the optional XML attribute ’id’ in the element INSDQualifier. The value of the ’id’ attribute must start with the letter ’q’ and continue with any positive integer. ’id’ values do not need to be numbered consecutively but an ’id’ attribute must be unique to one INSDQualifier element, i.e., the attribute value must only be used once in a sequence listing file.’；

(e) 更新第 77 段，在第一行‘organism’和‘for’之间加入空格和一个”：‘ ...and “organism″ for amino acid sequences...’；

(f) 更新第 79 段，示例 2：删除‘mol_type’前的空格；

(g) 将第 86 段更新为：‘The use of free text must be limited to a few short terms indispensable for the understanding of a characteristic of the sequence. For each qualifier **other than the ″translation″ qualifier**, the free text must not exceed 1000 characters.’；

(h) 将第 87(d)段的开头部分更新为‘For qualifiers with a language-dependent free text value, ...’；

(i) CWS 第九届会议批准更新以扩大《产权组织手册》的范围后，将所有‘工业产权’的提法更新为‘知识产权’。

8. 对产权组织标准 ST.26 附件的拟议修改可总结如下：

(a) 附件一，第六部分：更新表 5，将 frequency（频率）作为具有语种相关自由文本值的限定符；

(b) 附件一，第 6.21 部分：更新必需值格式，以表明频率限定符与语种相关；

(c) 附件一，第六部分：将表 5 的标题更新为‘List of qualifiers values for nucleotide sequences with language dependent free text values’；

(d) 附件一，第六和第八部分：更新注释以包含‘国际’，目前只提及‘国家或区域程序’；

(e) 附件一，第 6.77 部分：去除示例中在反密码子值之后但在小于号之前的空格；

(f) 附件一，第八部分：更新注释，使之与第六部分中的注释相一致。更新后的注释将以以下内容开始：‘Any qualifier value provided for a qualifier with a language dependent “free text” value format may require translation...’；

(g) 附件二：在 DTD 中加入以下评论，以明确应在 fileName 属性中提供的值：‘By default this will be set to the value provided for the project name in WIPO Sequence. If the value is identical to the actual ST.26 XML filename, it should be noted that Offices may enforce their requirements for the filename used which may restrict which characters are allowable for submitted electronic files. Please refer to the WIPO Sequence and ST.26 Knowledge Base for further details on Offices’ naming conventions for electronic files’；

(h) 附件六，示例 3(c)-1，问题 3：将术语‘酰化’改成‘乙酰化’；

(i) 附件六，示例 3(k)-2，问题 3：将第三段中提到的对序列编号 12 的引用改为：‘the feature key “misc_difference” with feature location “10” should be used together with two “replace” qualifiers where the value for one would be ″~~g~~k″ and the second would be ″c″’；

(j) 附件六，示例 7(b)-4 和 7(b)-5：将序列编号 89、90 和 91 改为序列编号 90、91 和 92；

(k) 附件六和附件六的附录：加入三个新示例，即示例 94-2、示例 30-2 和示例 12-1。这些新示例作为本文件的附件一全部提供；

(l) 附件七，场景 4：将第一句话改为以下文本：‘For both nucleotide sequences and amino acid sequences, ST.26 has the mandatory feature key ″source with two mandatory qualifier.’；

(m) 附件七，场景 4：将核苷酸序列表的标题改为：‘ST.26 - feature key 5.37 source; mandatory qualifier 6.39 mol_type (see ST.26 paragraph 75)’；

(n) 附件七，场景 4：将氨基酸序列表的标题改为‘ST.26 - feature key 7.30 source; mandatory qualifier 8.1 mol_type (see ST.26 paragraph 75)’；

(o) 附件七，场景 8：更新所含表格的第 18 行，在第四栏增加以下文本：**‘TATA_signal and (if <223> present):’**，同时从注释中删除以下文本：‘~~(if <223> present)~~’；

(p) 附件六的附录：因更新附件六而需要作出的任何相应修改。附件六的附录作为本文件附件二提供。

关于更新 ST.26 版本的指导意见

9. 如上所述，在 CWS 第九届会议上,国际局注意到关于定期更新产权组织 ST.26 及其可能产生影响的关切（见本文件第 4 段）。

10. 应当指出的是，根据需要并经 CWS 批准，产权组织 ST.26 每年最多只更新一次。在提交给 CWS 第十届会议的拟议修订中，既有编辑上的也有实质性的修正。主体部分的编辑性修改应对用户（包括知识产权局在内）没有实质性影响。对附件六和附件七的实质性修改也不应对知识产权局和用户产生影响。

11. 国际局将在产权组织 ST.26 的未来版本中提供一个适当的版本号，即 X.Y，向该标准的用户说明这一修订版是否视为向后兼容前一版本。产权组织 ST.26 的小幅更新仅以第二位数字（‘Y’）的增量来表示，并被认为是向后兼容的。主要更新以第一位数字（‘X’）的增量来表示，可能对你的 IT 系统产生影响，可能需要设备更新。DTD 的版本号可能与主体及其他附件的版本号不同。

12. 对于产权组织 ST.26 的 DTD，该新版本除了增加一条评论外没有任何拟议修改，而对主体和附件的拟议修改大多数是编辑性的或者为了澄清，因此在这种情况下，DTD 的版本号将仍为第 1.3 版，而该标准的版本号将更新为第 1.6 版。

13. 如果拟议修订得到 CWS 批准，国际局将与序列表工作队合作，将批准的内容纳入产权组织标准 ST.26，并在 2022 年公布。

14. 请标准委员会：

(a) 注意本文件及本文件附件的内容；并

(b) 审议并批准上文第 7 和第 8 段以及本文件附件一和附件二所述的对产权组织标准 ST.26 的拟议修订。

[后接附件一（附件六的拟议新示例）]

PROPOSED NEW EXAMPLES FOR ANNEX VI OF STANDARD ST.26

Paragraph 94 – Variant sequence disclosed as a single sequence with enumerated alternative residues

Example 94-2 – Representation of single sequence with enumerated alternative amino acids that may be modified amino acids

A patent application describes the following polypeptide:

Leu-Glu-Tyr-Cys-Leu-Lys-Arg-Trp-Xaa-Glu-Thr-Ile-Ser-His-Cys-Ala-Trp

where Xaa can be Ile, Ala, Phe, Tyr, aIle, MeIle, or Nle.

Question 1: Does ST.26 require inclusion of the sequence(s)?

YES

The enumerated peptide provides 16 specifically defined amino acids. Therefore, the sequence must be included in a sequence listing as required by ST.26 paragraph (7)(b).

Question 3: How should the sequence(s) be represented in the sequence listing?

The most restrictive ambiguity symbol that can encompass “Ile, Ala, Phe, Tyr, aIle, MeIle, or Nle” is “X”. Therefore, the sequence must be included in a sequence listing as:

LEYCLKRWXETISHCAW (SEQ ID NO: xx)

ST.26 paragraph 30 requires that “[a] modified amino acid must be further described in the feature table”. However, paragraph 30 does not require any specific feature key be used to describe modified amino acids. While paragraph 30 describes the use of feature keys “CARBOHYD”, “LIPID”, “MOD_RES”, and “SITE”, these feature keys are more appropriate for scenarios where the modified amino acid is not within a list of alternatives for a specific location. In this example, the feature key “VARIANT” satisfies the requirement of paragraph 30 since it allows for the inclusion of all of the alternatives for the variant site. So, the feature key “VARIANT” with the qualifier “note” and a qualifier value “Ile, Ala, Phe, Tyr, aIle, MeIle, or Nle” should be used to describe the variant site at position 9. The use of a second feature key such as “SITE” with a qualifier “note” may be used to further identify the modified amino acids found at position 9.

Relevant ST.26 paragraph(s): 3(a), 7(b), 27, 30, **94**, 96, and Annex I, Section 4, Table 4

Paragraph 30 – Annotation of a modified amino acid

Example 30-2 – Post-translationally modified amino acids

A patent application describes the following polypeptide:

Leu-Glu-Tyr-Cys-Leu-Lys-**Arg**-Trp-Glu-Thr-Ile-Ser-His

wherein the Arg at position 7 may be post-translationally deiminated to citrulline.

Question 1: Does ST.26 require inclusion of the sequence(s)?

YES

The enumerated peptide provides 13 specifically defined amino acids. Therefore, the sequence must be included in a sequence listing as required by ST.26 paragraph (7)(b).

Question 3: How should the sequence(s) be represented in the sequence listing?

According to ST.26 paragraph 29, a modified amino acid should be represented in the sequence as the corresponding unmodified amino acid whenever possible.

Therefore, the sequence should be included in a sequence listing as:

LEYCLKRWETISH (SEQ ID NO: xx)

where the symbol “R” is used to represent the arginine at position 7.

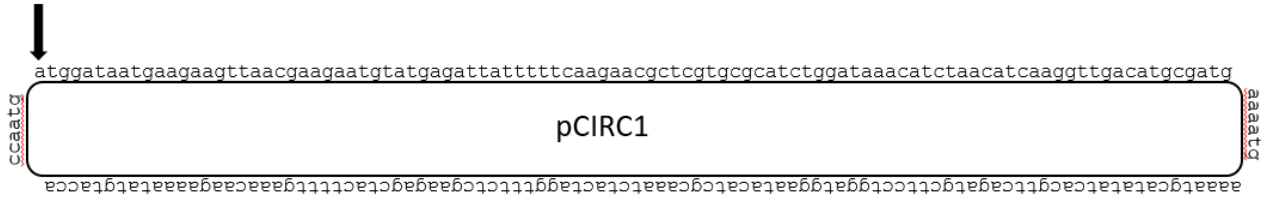
A further description indicating that the arginine at position 7 may be modified to citrulline is required. The modification of arginine to citrulline is a post-translational modification. Therefore, the feature key “MOD_RES” should be used together with the mandatory qualifier “note” to indicate that the arginine may be deiminated to form citrulline. The location descriptor in the feature location element is the residue position number of the modified arginine.

Relevant ST.26 paragraph(s): 3(a), 7(b), **30**, and Annex I, Section 7, Feature Key 7.18

Paragraph 12 – Circular nucleotide sequence

Example 12-1: Circular nucleotide sequence

A patent application contains the following figure, disclosing the DNA sequence of plasmid pCIRC1:



Question 1: Does ST.26 require inclusion of the sequence(s)?

YES

The enumerated nucleotide sequence has more than 10 specifically defined nucleotides. Therefore, the sequence must be included in a sequence listing as required by ST.26 paragraph (7)(a).

Question 3: How should the sequence(s) be represented in the sequence listing?

According to ST.26 paragraph 12, when nucleotide sequences are circular in configuration, the applicant must choose the nucleotide in residue position number 1. For the purposes of this example, the “a” residue identified by the arrow in the figure will be used as position 1. However, any residue may be chosen as position 1. With the residue indicated by the arrow as position 1, the sequence should be included in a sequence listing as:

atggataatgaagaagttaacgaagaatgtatgagattatttttcaagaacgctcgtgcgcatctggat
aaacatctaacatcaaggttgacatgcgatgaaaatgaaaatgcatatatcacgttcagatgcttcctg
gatggaatacatcgcaaatctactaggtttctcgaagagctacttttgaaacaagaaaatatgtaccac
caatg (SEQ ID NO: xx)

The sequence should be further described using feature key “misc_feature” with a location of “212^1”, which indicates that the last residue in the sequence, position 212, is linked to residue 1. A “note” qualifier must be included with a value indicating that the molecule is circular.

Relevant ST.26 paragraphs: 7(a), **12**, and Annex I, Section 5, Feature Key 5.15