

专利合作条约（PCT）工作组

第十八届会议

2025 年 2 月 18 日至 20 日，日内瓦

处理序列表

国际局编拟的文件

摘要

1. 产权组织标准 ST. 26 是 2022 年 7 月 1 日或之后提交的国际申请中核苷酸和氨基酸序列表必须使用的格式。在成功启用和过渡之后，该体系基本按预期高效运行，并在继续开展工作，进一步开发和改进处理该格式序列表的工具和程序。
2. 尽管如此，国际局仍希望简化处理过程，确保高质量的处理结果，满足专利申请人的要求，并适合国际单位、指定局和数据库提供方使用。例如，可以减少文件类型的数量。继续开展分析，研究改进验证安排的备选方案，并查明语言要求是否得到有效满足。产权组织标准委员会（标准委）序列表工作队正在考虑对该标准本身进行实质性修改，以扩大公开要求，并纳入短序列。

背景

3. 产权组织标准 ST. 26 是建议在专利申请中以 XML 格式列示核苷酸和氨基酸序列的标准，于 2022 年 7 月 1 日生效。对 PCT 实施细则和行政规程的修改（包括相关表格）的修改也同时生效，要求在该日期或之后提交的任何国际申请中，必须按照产权组织标准 ST. 26 进行序列公开。本文件介绍了最新发展情况，并提出了一些挑战和改进序列表处理现行做法的可能途径。

WIPO SEQUENCE 套件

4. 国际局开发了一个名为 WIPO Sequence 套件的软件套件，由以下两部分组成：
 - (a) WIPO Sequence: 一个独立的桌面应用程序，可用于 Windows、Linux 和 MacOS，协助申请人按照产权组织标准 ST. 26 编制和验证序列表；以及
 - (b) WIPO Sequence Validator: 一项在专利局环境内部运行的网络服务，用于检查提交的序列表是否符合产权组织标准 ST. 26。
5. WIPO Sequence 套件的开发和测试将优先进行，并经过与序列表工作队的协调。
6. 最新的 WIPO Sequence Validator 3.0.0 版于 2024 年 10 月发布，与此前的版本相比，性能有了显著提高，并删除了过时的依赖关系。操作手册也以 HTML 格式提供，有 10 种语言版本，为各局提供了更方便、更精简的资源，支持下载代码段，以帮助更快、更可靠地入门。自该版本发布以来，各局报告了一些错误，目前正在解决之中。补丁版本 3.0.1 计划于 2025 年 2 月发布。
7. WIPO Sequence 3.0.0 版仍处于测试阶段。同样，弃用的组件也已升级，WIPO Sequence Validator 现已纳入，为申请人及其代表确保证具有 100% 的一致性。

产权组织标准 ST. 26 的开发

8. 产权组织标准 ST. 26 的最新版本 1.7 于 2024 年 7 月 1 日生效，应当应用于该日期或之后提交的所有国际、国家和地区申请。与 1.6 版相比，修改主要是在指南文件中作了说明并新增了若干示例。序列表工作队目前正在审议两项实质性修订，供将于 2025 年 11 月召开的产权组织标准委员会第十三届会议审议。第一项修订由欧洲专利局提出，涉及取消不允许将短序列列入序列表的最低长度要求。另一项由美国专利商标局提出，要求在序列表中列入某些肽和核苷酸类似物。如果这些实质性修改获得批准，产权组织标准 ST. 26 的新版本将是 2.0 版。其中第一项修改还要求在新版标准生效日之前对 WIPO Sequence 套件进行重大更新。

优先权文件电子交换的数据包格式

9. 在 2024 年 9 月 16 日至 19 日召开的第十二届会议上，产权组织标准委员会通过了关于优先权文件电子交换数据包格式的新产权组织标准 ST. 92。该标准旨在改进优先权文件的处理，允许以产权组织标准 ST. 26 的 XML 格式，而不仅仅是 PDF 文件，进行结构化文本格式，特别是序列表的电子交换。会议初步商定，力争在 2027 年 7 月 1 日前实施新标准。这将涉及国际局为 PCT 体系和产权组织优先权文件数字查询服务（DAS）以及各国家局进行信息技术开发工作。

PCT 序列表的处理

10. 自 2022 年 7 月以来，国际局已收到约 32,000 件包含序列表的国际申请。对大多数申请而言，处理过程简单且符合预期，但也有相当数量的申请存在问题，需要工作人员提供进一步意见。

序列表验证

11. WIPO Sequence Validator 提供两种类型的验证：“形式验证”和“全面验证”。形式验证检查文件格式是否为 XML 格式以及是否符合产权组织标准 ST. 26 文档类型定义（DTD）。全面验证从形式验证检查开始，然后检查是否符合从产权组织标准 ST. 26 正文中衍生出来的一系列验证规

则。即使新版本的性能有所提高，WIPO Sequence Validator 操作手册仍建议仅将全面验证作为批处理的一部分使用，因为它比形式验证耗时长得多。

12. 使用 WIPO Sequence 编制的序列列表不应包含可通过 WIPO Sequence Validator 识别的错误（尽管可能会有警告），因为申请人生成序列列表时使用的是相同的验证规则。但是，在收到的大量序列列表中，虽然标明是由 WIPO Sequence 生成，但显然使用文本编辑器进行了一些手工后处理，其结果未经 WIPO Sequence 验证，以确保该序列列表仍然符合产权组织标准 ST. 26。

13. 目前，国际局的系统在将序列列表上传到 ePCT 时，或在通过 PCT-EDI 或其他传输方式从受理局和国际单位接收序列列表时，都没有使用 WIPO Sequence Validator 进行的全面验证检查；这些检查要由需要掌握这类详细信息国家局完成。此外，现行 PCT 受理局指南和行政规程（附件 C）并不要求受理局或国际局验证序列列表。但是，国际局正在评估 WIPO Sequence Validator 3.0.0 版，以确定是否可以通过 ePCT 和 PATENTSCOPE 提供全面验证，或是作为收到序列列表时执行的常规行动的结果，或根据要求提供。

根据细则 13 之三提交的序列列表

14. 产权组织标准 ST. 26 启用后，使用细则 13 之三，为国际检索的目的要求提交序列列表的情况有所减少，但没有预期的那么显著。原本的设想是，公开序列并被要求提交序列列表的国际申请都包含符合标准的序列列表，在提交前经过验证，并被认为适合国际检索。然而，约有 14% 包含序列列表的国际申请根据细则 13 之三提交了序列列表。它们包括：

- (a) 国际申请在提交时所包含的序列列表被认为有缺陷，需要提供符合标准的序列列表以供检索；以及
- (b) 最初没有提交序列列表，但国际申请公开了本应纳入序列列表的序列。

15. 下表 1 列出的是，截取 2024 年 12 月末这一时间节点，在 2022 年 7 月 1 日或此后提交并且国际局已收到某种序列列表的国际申请的数量，按受理局开列。第二至第四栏显示的是：

- (a) 包含任一类型序列列表的国际申请的数量；
- (b) 除了提交国际申请时已包含的序列列表，还包含细则 13 之三序列列表的国际申请的数量；以及
- (c) 提交国际申请未包含序列列表，但包含细则 13 之三序列列表的国际申请的数量。

表 1：按受理局开列的包含序列列表的国际申请数量

| 受理局 | 包含序列列表的全部国际申请 | 除提交申请时包含的序列列表，还包含细则 13 之三序列列表 | 仅包含后续提交的细则 13 之三序列列表 |
|-----|---------------|-------------------------------|----------------------|
| AU | 289 | | 13 |
| BR | 46 | | 1 |
| CA | 281 | 5 | 33 |
| CL | 6 | | |
| CN | 6,795 | 7 | 83 |
| CU | 12 | 9 | |
| CZ | 28 | | 4 |
| DE | 10 | | |
| DK | 7 | | 1 |

| 受理局 | 包含序列表的 全部国际申请 | 除提交申请时包含的序列表， 还包含细则 13 之三序列表 | 仅包含后续提交的细则 13 之三序列表 |
|-----|------------------|---------------------------------|------------------------|
| EP | 4,722 | 94 | 289 |
| ES | 60 | | 3 |
| FI | 30 | 1 | |
| FR | 46 | 2 | 1 |
| GB | 558 | 18 | 73 |
| GR | 4 | 1 | |
| HU | 13 | | 2 |
| IB | 1,235 | 66 | 129 |
| ID | 1 | 1 | |
| IL | 313 | 39 | 3 |
| IN | 89 | 4 | 10 |
| IT | 18 | 1 | 3 |
| JP | 1,882 | 1 | |
| KR | 2,237 | 1 | |
| MX | 3 | 1 | |
| MY | 5 | | |
| NL | 89 | | 46 |
| NO | 3 | | 1 |
| NZ | 23 | | 5 |
| PE | 1 | | |
| PH | 2 | | |
| PL | 43 | 6 | 6 |
| PT | 5 | 1 | |
| QA | 1 | | |
| RU | 44 | 3 | 1 |
| SE | 26 | 1 | 1 |
| SG | 190 | 1 | 7 |
| SK | 2 | | 1 |
| TH | 11 | 2 | 4 |
| TR | 43 | 4 | 10 |
| US | 12,179 | 991 | 2,336 |
| ZA | 1 | | |

16. 根据细则 13 之三提交序列表的原因需要进一步调查，但这其中既有国际检索单位的原因，也有序列表来源的原因。在一些国家，申请人在准备国际申请时会包含细则 5.2(a) 要求纳入序列表的序列，但却不提供序列表作为国际申请的一部分，这种情况比其他国家更常见。有些国际单位对提交的序列表的检查比其他国际单位更彻底，如果发现序列表有缺陷，或者申请正文中包含未纳入序列表的序列，这些国际单位更有可能根据细则 13 之三为国际检索的目的要求提供符合标准的序列表。

序列表处理过程中的人工检查

17. 欧洲专利局和美国专利商标局向序列表工作队提供了详细信息，涉及除了 WIPO Sequence 和 WIPO Sequence Validator 自动进行的检查，它们还对序列表进行哪些检查，这些检查触发了许多

提交细则 13 之三序列列表的要求。有关这些检查的信息可能对其他国家局改进对其收到的序列列表文件的处理很有价值。

18. 需要进一步分析常见的缺陷类型及其对国际检索的影响，以确定应考虑在以下方面采取何种行动（如果有的话）：

- (a) 对申请人进行教育，使其了解将序列列表作为说明书的一部分纳入申请的要求和产权组织标准 ST. 26 的要求；
- (b) WIPO Sequence 套件所需的改进；
- (c) 受理局、国际单位和国际局要进行的验证；
- (d) 关于在什么情况下为国际检索目的要求提交序列列表是有用的做法的指南（注意根据细则 13 之三的通知由单位酌情决定，细则 13 之三序列列表不成为国际申请的一部分，因此形式缺陷可能仍需要在第二章或国家阶段进行改正）。

序列列表相关文件类型

19. 下表 2 列出的是目前与序列列表有关的各类文件编码，以及自 2022 年 7 月以来提交的申请中每类文件的相应使用次数。由于国际局并不接收所有相关文件（尤其是根据细则 13 之三提交的文件），这份清单就 PCT 总体而言可能并不完整，但可作为很好的参考。申请人或形式审查员不易了解和选择不同的特殊情况所对应的正确文件类型。由于分不清多个相似的文件类型，或由于不小心点击了下拉列表中不正确的文件类型等错误，可能会出现人为错误。应进一步考虑是否应取消其中一些文件类型。例如，文件类型“序列列表 - 改正 - 细则 13 之三”、“序列列表 - 更正 - 细则 13 之三”和“由国际初步审查单位更正的根据细则 13 之三的序列列表（细则 91）”可能被认为没有必要，因为它本身不是国际申请的一部分。如果细则 13 之三序列列表因任何原因出错，可以直接用新的序列列表取而代之，而不必说明这一变动是由于改正还是更正。

表 2：序列列表文件类型的使用

| 文件类型 | 数量 |
|------------------------------|--------|
| 序列列表 | 28,804 |
| 供国际检索使用的序列列表（细则 13 之三） | 3,550 |
| 与供国际检索使用的序列列表相关的声明（细则 13 之三） | 3,319 |
| 最初提交的序列列表 | 253 |
| 序列列表 - 信息 | 213 |
| 序列列表 - 改正 | 128 |
| 据称的序列列表（不符合 ST. 26） | 65 |
| 序列列表 - 更正 | 63 |
| 序列列表 - 更正 - 细则 13 之三 | 50 |
| 供国际公布使用的序列列表译文 | 20 |
| 援引加入的序列列表（细则 20.6） | 19 |
| 供国际检索使用的序列列表译文 | 11 |
| 序列列表验证报告 | 8 |
| 序列列表 - 修改 - 条约第 34 条 | 6 |
| 有关可专利性的国际初步报告第 II 章的序列列表 | 5 |
| 序列列表 - 改正 - 细则 13 之三 | 3 |
| 序列列表 - 后续提交（细则 20.5 之二） | 2 |

| | |
|------------------------------------|---|
| 序列表 - 后续提交 (细则 20.5) | 2 |
| 序列表占位页面经核准副本 | 1 |
| 由国际初步审查单位更正的序列表 (细则 91) | 1 |
| 由国际初步审查单位更正的根据细则 13 之三的序列表 (细则 91) | 0 |
| 供补充国际检索使用的序列表译文 | 0 |

20. 目前, 一些受理局会检查序列表中的著录数据, 并根据细则 26 通知申请人提交改正。自 2022 年 7 月以来, 国际局收到了 128 份因申请人名称、优先权日、发明名称等缺陷而改正的序列表。为了减少此类改正的数量, 可以简化序列表中的著录数据, 或者受理局可以不检查这些著录数据。

序列表的译文

21. 产权组织标准 ST. 26 要求一个 XML 文件分两部分提供: 一般信息部分和序列数据部分。

一般信息部分的语言项目

22. 序列表的一般信息部分允许提供有关申请号、申请日、申请人文档编号、最早优先权申请、申请人名称、发明人姓名和发明名称的信息。申请人、发明人姓名和发明名称可以用多种语言或字符集提供。该信息部分提供的信息旨在帮助确认序列表是附于正确的国际申请。如果这些信息与请求书表格中的信息不同, 或者使用了不同的语言, 受理局可以提请申请人注意, 但不能要求申请人改正, 申请应根据请求书表格中的信息继续处理 (行政规程附件 C 第 28 段)。

序列数据部分中的语言相关自由文本

23. 序列表的序列数据部分依据的是数据库供应方使用的行业标准, 其中包括包含自由文本 (“限定符值”) 的元素。该行业标准要求这些值不使用任何语言, 或者使用英文。尽管一些国家局表示, 即使说明书的语言可能不是英文, 它们也可以接受所包含的语言相关自由文本为英文的序列表, 但许多主管局表示, 必须允许或要求其申请人用说明书的语言提供语言相关自由文本。因此, 产权组织标准 ST. 26 要求语言相关限定符值应包含以下一个或两个元素:

- (a) INSDQualifier_value - 英文限定符值; 以及
- (b) NonEnglishQualifier_value - 另一种语言的限定符值, 由序列表根元素中的属性 nonEnglishFreeTextLanguageCode 表示。

24. 根据细则 12.1(d), 受理局可以规定其接受的语言相关自由文本限定符值的语言, 这些语言可能不同于其接受的一般说明书语言。此外, 受理局可以准许申请人用一种以上的语言提交包含语言相关自由文本的序列表, 但不能要求申请人这样做。从技术上讲, 标准只允许在序列表中使用英文和一种其他语言。

25. 这一灵活性被认为非常重要, 可以确保向数据库提供的信息使用的是行业数据库格式 (对此行业标准特别要求在相关字段中使用英文文本), 申请人也可以满足主管局在提交申请、国际检索和多个进入国家阶段方面可能存在的不同语言要求。允许同时使用英文和第二种语言, 就不需要为提交申请和国际检索使用不同语言这种情况分别提供列表。

26. 一个在 WIPO Sequence 中创建的项目允许提供多种语言版本, 并允许按要求生成所需语言的序列表。因此, 应当可以制定一个项目, 并根据申请要求轻松导出适当的语言版本。例如, 可以为国际申请生成一个所包含语言相关自由文本为日文和英文的序列表; 然后, 在进入国家阶段

时，可以为不同的指定局生成实质上相同的序列列表，其中包含仅有英文的文本、中英文文本和韩文文本。

27. 在 WIPO Sequence 中，不同的语言值还可以导出为 XLIFF 文件，这是翻译人员使用的一种标准格式，用于标明源语言和目标语言。该文件可被读入标准翻译软件，填入适当的值，并将结果导入 WIPO Sequence，而无需手工逐一添加目标语言值。尚不确定申请人是否充分了解或使用这些功能。还注意到，序列列表包含语言相关自由文本限定符的 ID 属性，在导入 XLIFF 文件时可以匹配不同的语言值。然而，出现了一些与这些属性有关的问题并导致了错误。目前正在考虑是否应保留这些 ID 属性，还是即使在生成序列列表后也不向用户显示。

28. WIPO Sequence 并不会验证一个项目中是否指定了适当的语言。所作的唯一检查是 INSDQualifier_value 中的文本是否符合标准第 40(b)段，以及是否只包含基本（无重音）拉丁字符集中的字符。

29. 考虑到可以提交所包含的自由文本使用两种语言（英文和一种其他语言）的序列列表，而国际检索单位的审查员大多精通申请语言，无论该语言是否为公布语言，因此在国际阶段很少提交序列列表译文也许并不奇怪。自 2022 年 7 月以来，国际局仅收到 31 份注明是为公布或检索目的的序列列表译文的文件。然而，对序列列表样本的检查表明，语言选项并不总是按预期使用。在某些情况下（包括上文第 20 段所述的作为改正提供的文件），提供序列列表译文只是为了在一般信息部分提供相关语言，而这方面不应有此要求。此外，尚不确定限定符标签的内容是否总是使用正确的语言，如果标明同时使用英文和第二种语言，限定符文本是否等同。国际局希望对于在国际阶段对序列列表中语言选项的使用进行更系统的分析，以确定是否应采取与上文第 18 段所述类型相同的行动。欢迎指定局提供关于在进入国家阶段时使用产权组织标准 ST. 26 格式序列列表译文的信息。

公布序列列表

30. PATENTSCOPE 较长时间以来已提供了以类似于产权组织标准 ST. 26 显示格式的人性化格式查看序列列表的功能。PATENTSCOPE 更新于 2024 年 11 月部署，采用了新的显示引擎，以更加方便查看超大序列列表。转换为串流进行，而不是必须加载整个文件才能处理。大型序列列表在以前可能需要一分钟或更长时间才能显示，或者完全显示失败，而现在，序列列表的开头几乎会立即显示，如果没有完全显示，也可按要求加载其他部分。

31. 请工作组注意文件 PCT/WG/18/14 的内容，并为进一步审查和分析 PCT 序列列表的处理提出评论意见。

[文件完]