



WIPO

WORLD
INTELLECTUAL PROPERTY
ORGANIZATION

■ Стандарт ВОИС ST.26

ВВЕДЕНИЕ

Учебный вебинар

Аспекты, которые будут сегодня рассмотрены

- Что такое перечень последовательностей?
- Базы данных о последовательностях INSDC
- Зачем нужен новый стандарт?
- Преимущества стандарта ВОИС ST.26
- Различия между стандартами ST.25 и ST.26
- Переход на стандарт ST.26
- Основы XML
- Основы стандарта ST.26: элементы перечня последовательностей
- Содержание стандарта ВОИС ST.26
- Приложение WIPO Sequence: введение

Что такое перечень последовательностей?

Перечень последовательностей...

- Содержит последовательности нуклеотидов и/или аминокислот, раскрытые в патентной заявке, и составляет часть описания
- Содержит описательную информацию о каждой последовательности, называемую «аннотацией»
- Составлен согласно требованиям соответствующего стандарта ВОИС (ST.25 или ST.26)
- Позволяет искать данные об описанной в изобретении последовательности:
 - в базе данных ведомства ИС
 - в общедоступных для поиска базах данных (базах данных INSDC)

Справочная информация: INSDC

- INSDC — Международное сотрудничество баз данных о нуклеотидных последовательностях:
 - DDBJ: Банк данных о ДНК Японии
 - EMBL-EBI: Европейский институт биоинформатики
 - NCBI: Национальный центр биотехнологической информации (GenBank)

- Ведомства ИС, которые подают данные о последовательностях, содержащиеся в опубликованных/поданных заявках, в базы данных INSDC:
 - Европейское патентное ведомство
 - Японское патентное ведомство
 - Корейский институт интеллектуальной собственности
 - Ведомство по патентам и товарным знакам США

- Базы данных INSDC общедоступны для поиска

Зачем нужен новый стандарт?

- В настоящее время перечни последовательностей составляются в соответствии со стандартом ВОИС ST.25

Однако...

- Стандарт ST.25 не соответствует требованиям INSDC, из-за чего при внесении последовательностей в общедоступные базы данных теряются данные
- Правила стандарта ST.25 нечеткие, и ведомства ИС во всем мире по-разному толкуют и применяют эти правила
- Типы последовательностей, которые сегодня распространены, не охватываются стандартом ST.25 (нуклеотидные аналоги, D-аминокислоты, разветвленные последовательности) и поэтому отсутствуют в базах данных, доступных для публичного поиска
- Данные неструктурированы, и стандарт ST.25 трудно использовать для автоматизированной проверки и обмена данными

Преимущества стандарта ВОИС ST.26 (1)

- Обеспечивает принятие одного и того же перечня последовательностей по всему миру*
- Служит в качестве ориентира для согласованного применения ведомствами ИС правил, касающихся последовательностей
- Уточняет, какие последовательности необходимо или можно включать в перечень последовательностей, и как эти последовательности должны быть представлены
- Повышает качество подаваемых заявок благодаря XML-структуре перечней последовательностей
- Расширяет возможности для автоматической проверки данных и оптимизации процесса обработки заявок ведомствами ИС

* Не считая необходимости перевода зависящих от языка квалификаторов в свободном формате на язык подачи заявки для некоторых ведомств ИС, которые могут потребовать составления отдельных перечней последовательностей

Преимущества стандарта ВОИС ST.26 (2)

- Обеспечивает соответствие данных требованиям поставщиков данных INSDC (DDBJ, EBI и NCBI), и аннотации к последовательностям (ключи характеристик и квалификаторы) будут включены в базы данных, доступные для публичного поиска
- Обеспечивает стандартизацию:
 - аннотаций характеристик
 - указания местоположения характеристик
 - названий и значений квалификаторов
 - представления вариантов последовательностей
- Содержит требование о включении дополнительных типов последовательностей (нуклеотидных аналогов, D-аминокислот, разветвленных последовательностей), что означает доступность большего объема данных о последовательностях для поиска

Различия между ST.25 и ST.26

ST.25	ST.26
Формат ASCII .txt с числовыми идентификаторами	Формат XML с элементами и атрибутами
<p><u>Не требуется</u> включать:</p> <ul style="list-style-type: none"> - D-аминокислоты - Линейные части разветвленных последовательностей - Нуклеотидные аналоги 	<p><u>Требуется</u> включать:</p> <ul style="list-style-type: none"> - D-аминокислоты - Линейные части разветвленных последовательностей - Нуклеотидные аналоги
<p>Аннотации к последовательностям:</p> <ul style="list-style-type: none"> - Только ключи характеристик 	<p>Аннотации к последовательностям:</p> <ul style="list-style-type: none"> - Ключи характеристик и квалификаторы
<p><u>Можно</u> включать последовательности:</p> <ul style="list-style-type: none"> - с менее чем 10 специально определенными нуклеотидами - с менее чем 4 специально определенными аминокислотами 	<p><u>Нельзя</u> включать последовательности:</p> <ul style="list-style-type: none"> - с менее чем 10 специально определенными нуклеотидами - с менее чем 4 специально определенными аминокислотами

Различия между ST.25 и ST.26

...в части общей информации

ST.25	ST.26
Можно указать информацию обо ВСЕХ приоритетных заявках	Можно указать информацию ТОЛЬКО о самой ранней приоритетной заявке
Можно указать имена ВСЕХ заявителей и изобретателей	Можно указать ТОЛЬКО одного заявителя и (необязательно) ОДНОГО изобретателя
Можно указать одно название изобретения	Можно указать несколько названий изобретения, причем каждое — на другом языке
При указании имени заявителя/изобретателя и названия изобретения можно использовать только основную латиницу	При указании имени заявителя/изобретателя можно использовать любые допустимые символы Unicode, а также можно добавить перевод или транслитерацию основной латиницей

Различия между ST.25 и ST.26

...в части данных последовательности (1)

ST.25	ST.26
<p>Типы последовательностей — только DNA, RNA или PRT</p>	<p>Типы последовательностей — DNA, RNA или AA с обязательным квалификатором «mol_type» для дополнительного описания молекулы</p>
<p>Названия организмов:</p> <ul style="list-style-type: none"> - Род/вид на латыни - Название вируса - «artificial sequence» («искусственная последовательность») - «unknown» («неизвестный») 	<p>Названия организмов:</p> <ul style="list-style-type: none"> - Род/вид на латыни - Название вируса - «synthetic construct» («синтетическая конструкция») - «unidentified» («неидентифицированный»)
<p>«u» означает урацил в нуклеотидных последовательностях</p>	<p>«t» означает урацил в РНК-последовательностях и тимин в ДНК-последовательностях</p>
<p>Аминокислотные последовательности состоят из трехбуквенных сокращений</p>	<p>Аминокислотные последовательности состоят из однобуквенных сокращений</p>

Различия между ST.25 и ST.26

...в части данных последовательности (2)

ST.25	ST.26
Переменные «n» и «Хаа» должны иметь определение, указанное в характеристике	Для переменных «n» и «X» используется значение по умолчанию, без определения
Формат указания местоположения характеристики определен нечетко	Строго определенные форматы указания местоположения характеристики; можно использовать «<» «>» в последовательностях всех типов и «^», «join», «order» и «complement» в нуклеотидных последовательностях
Допускаются «смешанные» последовательности — нуклеотидные последовательности с трансляцией в аминокислотную последовательность	«Смешанные» последовательности не допускаются; нуклеотидные трансляции включаются только в квалификаторы «translation»

Переход на стандарт ВОИС ST.26

- На пятой сессии КСВ была согласована дата перехода — **1 января 2022 г.**, названная датой «**радикального перехода**»
- **Все** ведомства интеллектуальной собственности (ВИС) совершат переход одновременно на международном уровне (на уровне РСТ), на национальном уровне и на региональном уровне
- В качестве даты для определения того, применяется ли в отношении заявки стандарт ST.25 или стандарт ST.26, будет использоваться дата подачи международной заявки, а НЕ дата приоритета
- Примечание: стандарт ST.25 продолжит применяться в отношении заявок с датой подачи до 1 января 2022 г.

Стандарт ВОИС ST.26

...что необходимо включать?

- Нуклеотидные последовательности:
 - 10 или более «специально определенных» и «пронумерованных»* остатков
 - включая последовательности с нуклеотидными аналогами, такими как пептидно-нуклеиновые кислоты (ПНК) и гликоль-нуклеиновые кислоты (ГНК)
- Аминокислотные последовательности:
 - 4 или более «специально определенных» и «пронумерованных» остатка
 - включая последовательности с D-аминокислотами
 - в перечень последовательностей требуется включать линейные части разветвленной последовательности

Стандарт ВОИС ST.26

...что необходимо включить?

- Что такое «специально определенный» применительно к нуклеотиду или аминокислоте?
- «Специально определенный» означает любой из перечисленных в приложении I нуклеотид, отличный от нуклеотида, который представлен символом «n», и любую из перечисленных в приложении I аминокислоту, отличную от аминокислоты, представленной символом «X». (Стандарт ВОИС ST.26, п. 3(m))
- При определении соответствия требованию о минимальной длине учитываются только «специально определенные» остатки:
 - 10 или более специально определенных нуклеотидов; или
 - 4 или более специально определенных аминокислот

5'- anctggcaan – 3'

только 8 специально определенных нуклеотидов;
не подлежит включению в перечень последовательностей

5'- agctggcaat – 3'

10 специально определенных нуклеотидов;
подлежит включению в перечень последовательностей

Стандарт ВОИС ST.26: документ в формате XML

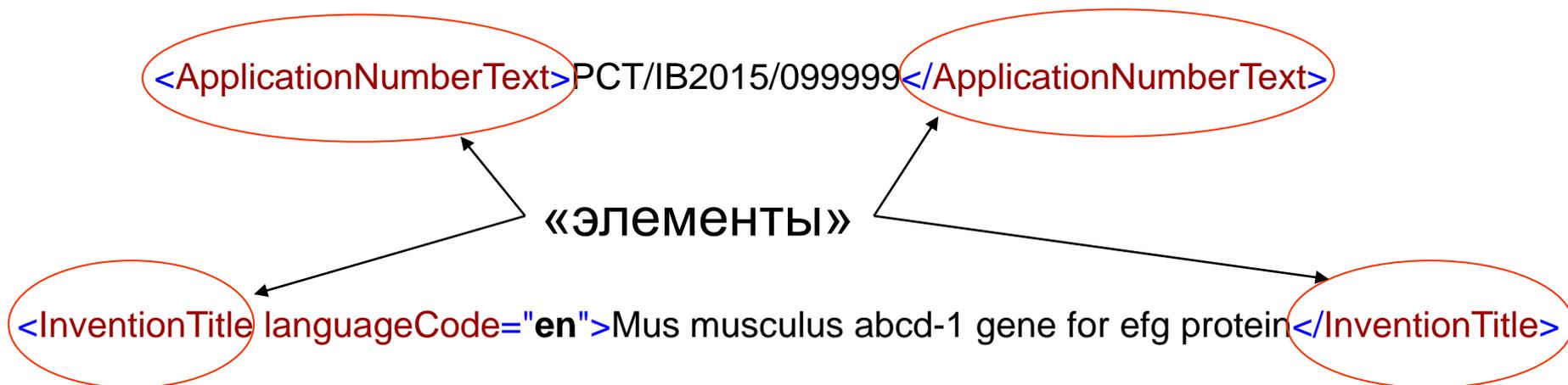
- Представляется в виде одного файла формата XML 1.0
- Должен соответствовать требованиям к определению типа документа (DTD), изложенным в стандарте ВОИС ST.26 (приложение II), и вытекающим из содержания стандарта общим правилам
- Должна использоваться кодировка Unicode UTF-8
- Структура перечня последовательностей:
 - Объявление XML
 - Объявление типа документа (DOCTYPE)
 - Корневой элемент
 - часть, содержащая общую информацию
 - часть, содержащая данные последовательности

ОСНОВЫ XML

- XML = eXtensible Markup Language (расширяемый язык разметки)
- Информация «размечается» с помощью описательных элементов и атрибутов
- Стандартизированное средство обмена данными, которые могут быть прочитаны как человеком, так и машиной
- DTD = Document Type Definition (определение типа документа) – задает структуру и допустимые элементы и атрибуты документа в формате XML

ОСНОВЫ XML

Элементы, атрибуты и значения (1)



ОСНОВЫ XML

Элементы, атрибуты и значения (2)

`<ApplicationNumberText>PCT/IB2015/099999</ApplicationNumberText>`

«ЗНАЧЕНИЯ ЭЛЕМЕНТОВ»

`<InventionTitle languageCode="en">Mus musculus abcd-1 gene for efg protein</InventionTitle>`

ОСНОВЫ XML

Элементы, атрибуты и значения (3)

```
<ApplicationNumberText>PCT/IB2015/099999</ApplicationNumberText>
```

«атрибут»

```
<InventionTitle languageCode="en">Mus musculus abcd-1 gene for efg protein</InventionTitle>
```

«значение атрибута»

ОСНОВЫ XML

Зарезервированные символы

При указании значения элемента должны быть заменены соответствующим сочетанием символов

Зарезервированный символ	Сочетание символов
<	<
>	>
&	&
“	"
'	'

Пример: желаемое местоположение характеристики — «<50..62»

`<INSDFeature_location><50..62</INSDFeature_location>` 

`<INSDFeature_location><50..62</INSDFeature_location>` 

Стандарт ВОИС ST.26: пример

```

<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE ST26SequenceListing PUBLIC "-//WIPO//DTD Sequence Listing 1.3//EN" "ST26SequenceListing_V1_3.dtd">
<ST26SequenceListing dtdVersion="V1_3" fileName="for WIPO Training" softwareName="WIPO Sequence" softwareVersion="1.1.0-beta4" productionDate="2021-03-24">
  <ApplicationIdentification>
    <IPOfficeCode>IB</IPOfficeCode>
    <ApplicationNumberText>PCT/IB2015/099999</ApplicationNumberText>
    <FilingDate>2015-01-31</FilingDate>
  </ApplicationIdentification>
  <ApplicantFileReference>ABC123</ApplicantFileReference>
  <EarliestPriorityApplicationIdentification>
    <IPOfficeCode>IB</IPOfficeCode>
    <ApplicationNumberText>PCT/IB2014/111111</ApplicationNumberText>
    <FilingDate>2014-01-31</FilingDate>
  </EarliestPriorityApplicationIdentification>
  <ApplicantName languageCode="en">Shutsugan Pharmaceuticals Kabushiki Kaisha</ApplicantName>
  <InventionTitle languageCode="en">Mus musculus abcd-1 gene for efg protein</InventionTitle>
  <SequenceTotalQuantity>1</SequenceTotalQuantity>
  <SequenceData sequenceIDNumber="1">
    <INSDSeq>
      <INSDSeq_length>52</INSDSeq_length>
      <INSDSeq_moltype>DNA</INSDSeq_moltype>
      <INSDSeq_division>PAT</INSDSeq_division>
      <INSDSeq_feature-table>
        <INSDFeature>
          <INSDFeature_key>source</INSDFeature_key>
          <INSDFeature_location>1..52</INSDFeature_location>
          <INSDFeature_qual>
            <INSDQualifier>
              <INSDQualifier_name>mol_type</INSDQualifier_name>
              <INSDQualifier_value>genomic DNA</INSDQualifier_value>
            </INSDQualifier>
            <INSDQualifier id="q2">
              <INSDQualifier_name>organism</INSDQualifier_name>
              <INSDQualifier_value>Mus musculus</INSDQualifier_value>
            </INSDQualifier>
          </INSDFeature_qual>
        </INSDFeature>
      </INSDSeq_feature-table>
      <INSDSeq_sequence>atgaaattaaaacataaaarggatgataaaatgagatttgatataaaaagg</INSDSeq_sequence>
    </INSDSeq>
  </SequenceData>
</ST26SequenceListing>

```

Стандарт ВОИС ST.26: компоненты (1)

Строка 1 – объявление XML
(ST.26, п. 39(a))

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE ST26SequenceListing PUBLIC "-//WIPO//DTD Sequence Listing 1.3//EN" "ST26SequenceListing_V1_3.dtd">
<ST26SequenceListing dtdVersion="V1_3" fileName="for WIPO Training" softwareName="WIPO Sequence"
softwareVersion="1.1.0-beta4" productionDate="2021-03-24">
  <ApplicationIdentification>
    <IPOfficeCode>IB</IPOfficeCode>
    <ApplicationNumberText>PCT/IB2015/099999</ApplicationNumberText>
    <FilingDate>2015-01-31</FilingDate>
  </ApplicationIdentification>
  <ApplicantFileReference>ABC123</ApplicantFileReference>
  <EarliestPriorityApplicationIdentification>
    <IPOfficeCode>IB</IPOfficeCode>
    <ApplicationNumberText>PCT/IB2014/111111</ApplicationNumberText>
    <FilingDate>2014-01-31</FilingDate>
  </EarliestPriorityApplicationIdentification>
  <ApplicantName languageCode="en">Shutsugan Pharmaceuticals Kabushiki Kaisha</ApplicantName>
  <InventionTitle languageCode="en">Mus musculus abcd-1 gene for efg protein</InventionTitle>
  <SequenceTotalQuantity>1</SequenceTotalQuantity>
  <SequenceData sequenceIDNumber="1">
```

Стандарт ВОИС ST.26: компоненты (2)

Строка 2 – объявление типа документа (DOCTYPE)

(ST.26, п. 39(b))

```

<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE ST26SequenceListing PUBLIC "-//WIPO//DTD Sequence Listing 1.3//EN" "ST26SequenceListing_V1_3.dtd">
<ST26SequenceListing dtdVersion="V1_3" fileName="for WIPO Training" softwareName="WIPO Sequence"
softwareVersion="1.1.0-beta4" productionDate="2021-03-24">
  <ApplicationIdentification>
    <IPOfficeCode>IB</IPOfficeCode>
    <ApplicationNumberText>PCT/IB2015/099999</ApplicationNumberText>
    <FilingDate>2015-01-31</FilingDate>
  </ApplicationIdentification>
  <ApplicantFileReference>ABC123</ApplicantFileReference>
  <EarliestPriorityApplicationIdentification>
    <IPOfficeCode>IB</IPOfficeCode>
    <ApplicationNumberText>PCT/IB2014/111111</ApplicationNumberText>
    <FilingDate>2014-01-31</FilingDate>
  </EarliestPriorityApplicationIdentification>
  <ApplicantName languageCode="en">Shutsugan Pharmaceuticals Kabushiki Kaisha</ApplicantName>
  <InventionTitle languageCode="en">Mus musculus abcd-1 gene for efg protein</InventionTitle>
  <SequenceTotalQuantity>1</SequenceTotalQuantity>
  <SequenceData sequenceIDNumber="1">

```

Стандарт ВОИС ST.26: компоненты (3)

Строка 3 – корневой элемент
(ST.26, п. 43)

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE ST26SequenceListing PUBLIC "-//WIPO//DTD Sequence Listing 1.3//EN" "ST26SequenceListing_V1_3.dtd">
<ST26SequenceListing dtdVersion="V1_3" fileName="for WIPO Training" softwareName="WIPO Sequence"
softwareVersion="1.1.0-beta4" productionDate="2021-03-24">
```

```
<ApplicationIdentification>
  <IPOfficeCode>IB</IPOfficeCode>
  <ApplicationNumberText>PCT/IB2015/099999</ApplicationNumberText>
  <FilingDate>2015-01-31</FilingDate>
</ApplicationIdentification>
<ApplicantFileReference>ABC123</ApplicantFileReference>
<EarliestPriorityApplicationIdentification>
  <IPOfficeCode>IB</IPOfficeCode>
  <ApplicationNumberText>PCT/IB2014/111111</ApplicationNumberText>
  <FilingDate>2014-01-31</FilingDate>
</EarliestPriorityApplicationIdentification>
<ApplicantName languageCode="en">Shutsugan Pharmaceuticals Kabushiki Kaisha</ApplicantName>
<InventionTitle languageCode="en">Mus musculus abcd-1 gene for efg protein</InventionTitle>
<SequenceTotalQuantity>1</SequenceTotalQuantity>
<SequenceData sequenceIDNumber="1">
```

Общая информация:
(ST.26, пп. 38(a),
45-49)

Стандарт ВОИС ST.26: компоненты (4)

```

<SequenceData sequenceIDNumber="1">
  <INSDSeq>
    <INSDSeq_length>52</INSDSeq_length>
    <INSDSeq_moltype>DNA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..52</INSDFeature_location>
        <INSDFeature_quals>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>genomic DNA</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q2">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>Mus musculus</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_quals>
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>atgaaattaaaacataaaaarggatgataaaatgagatttgatataaaaaagg</INSDSeq_sequence>
  </INSDSeq>
</SequenceData>
</ST26SequenceListing>

```

Данные последовательности

(ST.26, пп. 38(b), 50-100)

Стандарт ВОИС ST.26: общая информация (1)

- Раздел «Application Identification» («Идентификация заявки»)
 - обязательно указываются, номер заявки, дата подачи и код ведомства ИС, если они известны
 - если вышеуказанные сведения неизвестны, достаточно указать номер дела заявителя

```
<ApplicationIdentification>
  <IPOfficeCode>IB</IPOfficeCode>
  <ApplicationNumberText>PCT/IB2015/099999</ApplicationNumberText>
  <FilingDate>2015-01-31</FilingDate>
</ApplicationIdentification>
<ApplicantFileReference>ABC123</ApplicantFileReference>
<EarliestPriorityApplicationIdentification>
  <IPOfficeCode>IB</IPOfficeCode>
  <ApplicationNumberText>PCT/IB2014/111111</ApplicationNumberText>
  <FilingDate>2014-01-31</FilingDate>
</EarliestPriorityApplicationIdentification>
```

Стандарт ВОИС ST.26: общая информация (2)

- Раздел «Priority Application» («Приоритетная заявка»)
 - в перечне последовательностей можно указать только одну приоритетную заявку, и это должна быть самая ранняя приоритетная заявка
 - этот раздел заполнять обязательно, если испрашивается приоритет

```
<ApplicationIdentification>
  <IPOfficeCode>IB</IPOfficeCode>
  <ApplicationNumberText>PCT/IB2015/099999</ApplicationNumberText>
  <FilingDate>2015-01-31</FilingDate>
</ApplicationIdentification>
<ApplicantFileReference>ABC123</ApplicantFileReference>
<EarliestPriorityApplicationIdentification>
  <IPOfficeCode>IB</IPOfficeCode>
  <ApplicationNumberText>PCT/IB2014/111111</ApplicationNumberText>
  <FilingDate>2014-01-31</FilingDate>
</EarliestPriorityApplicationIdentification>
```

Стандарт ВОИС ST.26: общая информация (3)

- Раздел «Applicant and Inventor Name» («Имя заявителя и изобретателя»)
 - в перечне последовательностей можно указать только одно имя заявителя и одно имя изобретателя, и это должны быть «основные» заявитель и изобретатель
 - имя заявителя указывать обязательно; имя изобретателя указывается по желанию
 - код языка для имени заявителя и имени изобретателя указывать обязательно
 - если имя заявителя и/или имя изобретателя содержит какие-либо символы, помимо основной латиницы в кодировке Unicode, необходимо указать транслитерацию или перевод с использованием основной латиницы

```
<ApplicantName languageCode="ja">出願製薬株式会社</ApplicantName>
<ApplicantNameLatin>Shutsugan Pharmaceuticals Kabushiki Kaisha</ApplicantNameLatin>
<InventorName languageCode="ja">特許 太郎</InventorName>
<InventorNameLatin>Taro Tokkyo</InventorNameLatin>
```

Стандарт ВОИС ST.26: общая информация (4)

- Раздел «Invention Title» («Название изобретения»)
 - необходимо указать минимум одно название изобретения на языке подачи
 - можно указать дополнительные названия на других языках
 - для каждого названия обязательно указывается код языка

```
<InventionTitle languageCode="en">My spectacular invention</InventionTitle>  
<InventionTitle languageCode="de">Meine spektakuläre Erfindung</InventionTitle>  
<InventionTitle languageCode="lv">Mans iespaidīgais izgudrojums</InventionTitle>  
<InventionTitle languageCode="ru">Мое зрелищное изобретение</InventionTitle>
```

Стандарт ВОИС ST.26: общая информация (5)

- Элемент «Sequence Total Quantity» («Общее количество последовательностей»)
 - обязательный элемент
 - общее количество последовательностей указывается с учетом пропущенных последовательностей

`<SequenceTotalQuantity>6</SequenceTotalQuantity>`

Стандарт ВОИС ST.26: последовательности (1)

- Обязательные элементы

```

<SequenceData sequenceIDNumber="1">
  <INSDSeq>
    <INSDSeq_length>52</INSDSeq_length>
    <INSDSeq_moltype>DNA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..52</INSDFeature_location>
        <INSDFeature_qual>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>genomic DNA</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q2">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>Mus musculus</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_qual>
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>atgaaattaaaacataaaaaggatgataaaatgagatttgatataaaaaagg</INSDSeq_sequence>
  </INSDSeq>
</SequenceData>

```

«Идентификационный номер
последовательности»
(«SEQ ID NO»)

Длина
последовательности

Тип молекулы
(DNA, RNA или AA)

Указание того,
к чему относится
последовательность
(значение
всегда «PAT»)

Стандарт ВОИС ST.26: последовательности (2)

- Обязательная характеристика «source» или «SOURCE»

```

<SequenceData sequenceIDNumber="1">
  <INSDSeq>
    <INSDSeq_length>52</INSDSeq_length>
    <INSDSeq_moltype>DNA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..52</INSDFeature_location>
        <INSDFeature_qual>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>genomic DNA</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q2">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>Mus musculus</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_qual>
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>atgaaattaaaacataaaaaggatgataaaatgagatttgatataaaaaagg</INSDSeq_sequence>
  </INSDSeq>
</SequenceData>

```

Для каждой последовательности требуется указать только одну характеристику «source»

Местоположение характеристики «source» должно охватывать всю последовательность

Два обязательных квалификатора: «mol_type» и «organism»

Стандарт ВОИС ST.26: последовательности (3)

Данные последовательности в формате ST.26:

```
<SequenceData sequenceIDNumber="1">
  <INSDSeq>
    <INSDSeq_length>52</INSDSeq_length>
    <INSDSeq_moltype>DNA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..52</INSDFeature_location>
        <INSDFeature_qual>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>genomic DNA</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q2">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>Mus musculus</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_qual>
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>atgaaattaaacataaaarggatgataaaatgagatttgatataaaaaagg</INSDSeq_sequence>
  </INSDSeq>
</SequenceData>
```

Стандарт ВОИС ST.26: последовательности (4)

- Обязательные квалификаторы «mol_type» и «MOL_TYPE»:
возможные значения

DNA

genomic DNA

other DNA

unassigned DNA

RNA

genomic RNA

mRNA

tRNA

rRNA

other RNA

transcribed RNA

viral cRNA

unassigned RNA

AA

protein

Стандарт ВОИС ST.26: последовательности (5)

- Возможные значения обязательных квалификаторов «organism» и «ORGANISM»:
 - название рода и вида на латыни, например «Mus musculus»
 - название рода с сокращением «sp.», например «Mus sp.»
 - название вируса, например «Torque teno virus 1»
 - «unidentified»
 - «synthetic construct»
- Общепринятые названия, например «mouse», нельзя использовать в качестве названия организма. При желании общепринятое название можно указать в перечне последовательностей в квалификаторе «note»

Стандарт ВОИС ST.26: последовательности (6)

Характеристики и квалификаторы

Помимо обязательной характеристики «source» или «SOURCE», заявители могут добавить различные необязательные характеристики для более подробного описания последовательности:

- для нуклеотидных последовательностей и аминокислотных последовательностей используются разные ключи характеристик
- каждая характеристика может иметь один или более необязательных квалификаторов, а также обязательный квалификатор

5.12. Feature Key	misc_binding
Definition	site in nucleic acid which covalently or non-covalently binds another moiety that cannot be described by any other binding key (primer_bind or protein_bind)
Mandatory qualifiers	bound_moiety
Optional qualifiers	allele function gene gene_synonym map note
Comment	note that the regulatory feature key and regulatory_class qualifier with the value "ribosome_binding_site" must be used for describing ribosome binding sites

(Стандарт ВОИС ST.26, приложение I, разделы 5-8)

Стандарт ВОИС ST.26: последовательности (7)

- Нуклеотидные последовательности:
 - все символы — в нижнем регистре
 - без пробелов, без нумерации
 - не содержат символа «u»; «t» означает урацил в РНК
 - «n» имеет значение по умолчанию «любой из ‘a’, ‘c’, ‘g’ или ‘t/u’»

Символ	Нуклеотид
a	аденин
c	цитозин
g	гуанин
t	тимин в ДНК/урацил в РНК (t/u)
m	a или c
r	a или g
w	a или t/u
s	c или g
y	c или t/u
k	g или t/u
v	a или c или g; не t/u
h	a или c или t/u; не g
d	a или g или t/u; не c
b	c или g или t/u; не a
n	a или c или g или t/u; «unknown» («неизвестная») или «other» («другая»)

Стандарт ВОИС ST.26: последовательности (8)

- Аминокислотные последовательности:
 - однобуквенные символы в верхнем регистре
 - без пробелов, без нумерации
 - «X» имеет значение по умолчанию «любой из 'A', 'R', 'N', 'D', 'C', 'Q', 'E', 'G', 'H', 'I', 'L', 'K', 'M', 'F', 'P', 'O', 'S', 'U', 'T', 'W', 'Y' или 'V'»

Символ	Аминокислота
A	Аланин
R	Аргинин
N	Аспарагин
D	Аспарагиновая кислота (аспартат)
C	Цистеин
Q	Глутамин
E	Глутаминовая кислота (глутамат)
G	Глицин
H	Гистидин
I	Изолейцин
L	Лейцин
K	Лизин
M	Метионин
F	Фенилаланин
P	Пролин
O	Пирролизин
S	Серин
U	Селеноцистеин
T	Треонин
W	Триптофан
Y	Тирозин
V	Валин
B	Аспарагиновая кислота или аспарагин
Z	Глутамин или глутаминовая кислота
J	Лейцин или изолейцин
X	A или R или N или D или C или Q или E или G или H или I или L или K или M или F или P или O или S или U или T или W или Y или V; «unknown» («неизвестная») или «other» («другая»)

Стандарт ВОИС ST.26: последовательности (8)

- Пропущенные последовательности: позволяют заявителю удалять данные последовательности из перечня последовательностей без необходимости изменения нумерации последующих последовательностей
 - INSDSeq_length, INSDSeq_moltype, INSDSeq_division присутствуют, но не имеют значений
 - Таблица характеристик отсутствует, характеристика «source» отсутствует
 - Элемент последовательности должен иметь значение «000»

```
<SequenceData sequenceIDNumber="7">  
  <INSDSeq>  
    <INSDSeq_length/>  
    <INSDSeq_moltype/>  
    <INSDSeq_division/>  
    <INSDSeq_sequence>000</INSDSeq_sequence>  
  </INSDSeq>  
</SequenceData>
```

Стандарт ВОИС ST.26: содержание

- **Основной текст** – требования к включению/представлению данных
- **Приложение I** – Контролируемая лексика на основе контролируемой лексики INSDC
- **Приложение II** – Определение типа документа согласно стандарту ST.26 (DTD)
- **Приложение III** – Образец перечня последовательностей стандарта ST.26 (файл XML)
- **Приложение IV** – Подмножество символов основной латиницы для использования в перечнях последовательностей стандарта ST.26 в формате XML
- **Приложение V** – Требования для обмена данными с использованием элемента INDS (только для ВИС)
- **Приложение VI** – Руководство с примерами
- **Дополнение к приложению VI** – Файл XML с данными всех последовательностей, приведенных в качестве примера в приложении VI
- **Приложение VII** – Рекомендации по преобразованию перечня последовательностей из перечня стандарта ST.25 в перечень стандарта ST.26

Стандарт ВОИС ST.26:

Основной текст

Номер пункта	Содержание
1-9	Введение; определения; сфера применения; отсылки
10-37	Представление последовательностей
38-49	Структура перечня последовательностей в формате XML
50-71	Часть с данными последовательности; таблица характеристик; ключи характеристик; обязательные ключи характеристик; местоположение характеристики
72-84	Квалификаторы характеристик; обязательные квалификаторы характеристик
85-100	Произвольный текст; кодирующие последовательности; варианты

Стандарт ВОИС.26, приложение I:

Контролируемая лексика

Раздел	Содержание
1	Перечень нуклеотидов (однобуквенные символы в нижнем регистре)
2	Перечень модифицированных нуклеотидов
3	Перечень аминокислот (однобуквенные символы в верхнем регистре)
4	Перечень модифицированных аминокислот
5	Ключи характеристик для нуклеотидных последовательностей
6	Квалификаторы для нуклеотидных последовательностей
7	Ключи характеристик для аминокислотных последовательностей (адаптировано из базы данных UniProt)
8	Квалификаторы для аминокислотных последовательностей
9	Таблицы генетических кодов

Стандарт ВОИС.26, приложение II:

Определение типа документа (DTD) согласно стандарту ST.26

- Сведения об определении типа документа (DTD) согласно стандарту ST.26: текущая версия 1.3

- Часть с общей информацией
 - Элементы, относящиеся к информации о патентной заявке

- Часть с данными последовательности
 - Подмножество INSDC DTD
 - Один или более элементов данных последовательности, при этом каждый элемент содержит информацию об одной последовательности

Стандарт ВОИС.26, приложение VI: Руководство

- Содержит 49 реальных примеров последовательностей и объяснение того, как правила стандарта ST.26 применяются к каждому из примеров.

- В каждом примере указано:
 1. Требуется ли, разрешено ли или запрещено ли включение последовательности в перечень последовательностей;
 2. Если включение последовательности в перечень последовательностей требуется или разрешено, то как эта последовательность должна быть представлена.

- Дополнение к приложению VI представляет собой перечень последовательностей стандарта ST.26 в формате XML, содержащий все примеры, представленные в руководстве.

Стандарт ВОИС.26, приложение VII:

Рекомендации по преобразованию перечня последовательностей из перечня стандарта ST.25 в перечень стандарта ST.26

- ...или как избежать добавления нового объекта (признака)
- Требования стандарта ST.26 отличаются от стандарта ST.25: стандарт ST.26 требует указания информации, указание которой не требуется стандартом ST.25
- Преобразование перечня последовательностей стандарта ST.25 в перечень последовательностей стандарта ST.26 всегда требует ввода данных заявителем
- Преобразование перечня последовательностей стандарта ST.25 в перечень последовательностей стандарта ST.26 не приведет к созданию нового объекта (признака), если следовать рекомендациям, изложенным в приложении VII
- Рассматривается 20 примеров преобразования с рекомендациями и примерами

Приложение WIPO Sequence (1)

- Компьютерное приложение, разработанное ВОИС для оказания содействия в подготовке, проверке и генерировании перечней последовательностей стандарта ST.26
- Государства-члены обратились к ВОИС с просьбой разработать этот общий инструмент для всех ведомств и заявителей на международном, национальном и региональном уровнях
- Использование WIPO Sequence с удобным пользовательским интерфейсом упрощает создание перечней последовательностей стандарта ST.26 в формате XML: нет необходимости редактировать непосредственно файл XML
- Загрузить последнюю версию можно по ссылке:
<https://www.wipo.int/standards/ru/sequence/index.html>

Приложение WIPO Sequence (2)

- Информацию о последовательности можно сохранить в проекте и проверить, а затем можно сгенерировать перечень последовательностей стандарта ST.26
- Можно импортировать данные из: перечней последовательностей стандарта ST.26, проектов стандарта ST.26, перечней последовательностей стандарта ST.25, файлов с множественными последовательностями (в формате multi-sequence), файлов с необработанными данными (в формате raw) и файлов в формате FASTA
- Можно выполнять проверку перечней последовательностей в формате XML
- Соответствующие характеристики, квалификаторы и названия организмов можно выбирать из раскрывающихся списков
- Сведения о заявителях и изобретателях можно хранить в базе данных «Лица и организации»
- Поддерживает экспорт и импорт файлов в формате XLIFF, используемом переводчиками

Приложение WIPO Sequence: страница проектов

Project name	Applicant file reference	Applicant name	Invention title	Status	Creation Date
16079428_no_source_sequence	H0075.70243US00	Ferring B.V.	STABLE LIQUID GONADOTROPIN FORMULATION	invalid	2021-02-01
All features and qualifiers	123abc	Simple Healthkit, Inc.	Compositions and Methods for Treating Cancer	modified	2021-02-01
Ark three letter aa code import from ST.25	abc123	Joe, Smith	Improper 3-letter AA code for import	new	2021-02-26
Beta 4 Test Project	abc123	Stephenson, Adam	Beta 4 Test Project	modified	2021-03-11
DNA RNA with u and t	uspto	Biotech, Inc.	Beta 4 test import ST25	modified	2021-03-15

Приложение WIPO Sequence: подробные сведения о проекте

Раздел «General Information» («Общая информация»)

WIPO | Sequence
 FOR WIPO TRAINING
VERIFICATION REPORT
FREE TEXT QUALIFIERS
IMPORT REPORT
DISPLAY THE SEQUENCE LISTING
HELP
PREFERENCES ENGLISH
Return to project home

GENERAL INFORMATION SEQUENCES

GENERAL INFORMATION

APPLICATION IDENTIFICATION

Application Identified Before the assignment of the application number Application filed Applicant file reference ABC123	IP Office IB - International Bureau of the World Intellectual Property Organization [WIPO] Application number PCT/IB2015/099999 Filing date 2015-01-30	
--	--	--

PRIORITY IDENTIFICATION

IP Office	Application Number	Filing date	Selected Earliest Priority Application
IB - International Bureau of the World Intellectual Property Organization [WIPO]	PCT/IB2014/111111	2014-01-30	Yes

APPLICANT & INVENTOR

Applicant name Shutsugan Pharmaceuticals Kabushiki Kaisha	Primary applicant	
---	-------------------	--

INVENTION TITLE

Invention title Mus musculus abcd-1 gene for efg protein	Language en - English	
--	-----------------------	--

Приложение WIPO Sequence: подробные сведения о проекте

Раздел «Sequences» («Последовательности»)

The screenshot displays the WIPO Sequence application interface. At the top, there is a navigation bar with the WIPO logo and the text 'WIPO | Sequence'. To the right of the logo are several menu items: 'FOR WIPO TRAINING', 'VERIFICATION REPORT', 'FREE TEXT QUALIFIERS', 'IMPORT REPORT', 'DISPLAY THE SEQUENCE LISTING', and 'HELP'. Further right are 'PREFERENCES' and 'ENGLI' with a dropdown arrow. On the far right is a 'Return to project home' button.

Below the navigation bar, there are two tabs: 'GENERAL INFORMATION' and 'SEQUENCES', with 'SEQUENCES' being the active tab. The main content area is titled 'SEQUENCE 1'.

Under 'SEQUENCE 1', there is a summary box containing the following information:

- Sequence Number [ID] 1
- Sequence Name test
- Length 52
- Molecule Type DNA
- Organism Mus musculus

To the right of this summary box is an edit icon (a pencil).

Below the summary box is a section titled 'FEATURES' with a dropdown arrow. Underneath is an 'Add feature' button. Below that is a table with the following data:

Feature Key	Location	Qualifiers
source	1..52	mol_type = genomic DNA organism = Mus musculus

Below the table is another section titled 'SEQUENCE' with a dropdown arrow. Underneath is a text area containing the sequence: 'atgaaattaa aacataaaar ggatgataaa atgagatttg atataaaaaa gg'. To the right of the sequence is the number '52' and an edit icon (a pencil).

At the bottom right of the sequence area, there are navigation arrows and the text '< 1/1 >'. Below the sequence area is an upward-pointing arrow.

Что дальше?

standards@wipo.int

Вопросы и ответы

Глоссарий: сокращения

- КСВ — Комитет по стандартам ВОИС
- DDBJ — Банк данных о ДНК Японии
- EMBL-EBI — Европейский институт биоинформатики
- ЕПВ — Европейское патентное ведомство
- INSDC — Международное сотрудничество баз данных о нуклеотидных последовательностях
- ВИС — ведомство интеллектуальной собственности
- NCBI — Национальный центр биотехнологической информации
- ВОИС — Всемирная организация интеллектуальной собственности