

# ■ WIPO 標準 ST.26

## 入門編

ウェビナートレーニング

# 本日のトピック

- 配列表とは何か？
- INSDC 配列データベース
- なぜ新しい標準が必要なのか
- WIPO ST.26 のメリット
- ST.25とST.26 の比較 – 両者の違いについて
- ST.26 への移行
- XML の基礎知識
- ST.26 の基礎知識 – 配列表の各部について
- WIPO ST.26 の内容
- WIPO Sequence の紹介

# 配列表とは何か

配列表とは...

- 特許出願に開示されている塩基配列および／またはアミノ酸配列を含み、明細書の一部を形成している
- アノテーションと呼ばれる各配列に関する記述的情報を含んでいる
- 関連するWIPO標準（ST.25またはST.26）に準拠している
- 発明の配列データを検索可能にする：
  - 知財庁内
  - 一般公開されているデータベース（INSDC データベース）

# 背景 - INSDC

- INSDC: 国際塩基配列データベース連携：
  - DDBJ: 日本DNAデータバンク
  - EMBL-EBI: 欧州バイオインフォマティクス研究所
  - NCBI: アメリカ国立生物工学情報センター (GenBank)
  
- INSDC データベースに、公開・発行済みの出願の配列データを提出する知財庁：
  - 欧州特許庁
  - 日本特許庁
  - 韓国特許庁
  - 米国特許商標庁
  
- INSDCデータベースは広く一般に検索可能

# なぜ新しい標準が必要なのか

- 現在、配列表は WIPO ST.25 に準拠して提出されている  
しかしながら...
- ST.25形式は INSDCの要件には準拠しておらず、一般に検索可能なデータベースに入力される際にデータが失われている
- ST.25の規則は明確でなく、世界中の知財庁が規則を独自に解釈し、施行している
- ST.25規則では近年一般的となった配列型（ヌクレオチド類似体、D-アミノ酸、分岐配列 など）の取り扱いがなく、それらの配列型が検索可能なデータベースに存在しない
- データが構造化されていない – ST.25形式を自動検証とデータ交換に使用することが困難である

# WIPO ST.26 のメリット (1)

- 単一配列表 の世界的な承認\*
- 当標準は配列に関する規則の適用について知財庁間の合意を確保するためのガイダンスとして機能する
- 配列表に記載する必要がある又は記載が許されている配列の開示内容および表現方法が明確になる
- XML配列表の構造により提出物の品質が向上する
- データ検証の自動化および知財庁による処理の合理化を促す

\* 言語依存のフリーテキストQualifierを特定の知財庁の出願言語に翻訳する必要がある場合があり、代替の配列表が必要になる場合がある

## WIPO ST.26 のメリット (2)

- INSDCデータベースプロバイダー（DDBJ、EBI、NCBI）の要件とのデータの互換性 - 配列のアノテーション（Feature keysおよびQualifiers）は広く一般に検索可能なデータベースに含まれる
  
- 標準化：
  - Feature annotations (配列の特徴に関する注釈)
  - Feature locations (配列の特徴の存在位置)
  - Qualifiers及びQualifiers の値
  - 配列変異型の表示
  
- 追加の配列型（ヌクレオチド類似体、D-アミノ酸、分岐配列）の記載を要件とすることで、より多くの配列データが検索可能になる

# WIPO ST.25とST.26の比較

ST.25	ST.26
数値識別子付きのASCII .txt	要素と属性を持つXML
<p><u>記載の必要なし</u>：</p> <ul style="list-style-type: none"> <li>- D-アミノ酸</li> <li>- 分岐配列の線形部分</li> <li>- ヌクレオチド類似体</li> </ul>	<p><u>記載の必要あり</u>：</p> <ul style="list-style-type: none"> <li>- D-アミノ酸</li> <li>- 分岐配列の線形部分</li> <li>- ヌクレオチド類似体</li> </ul>
<p>配列のアノテーション：</p> <ul style="list-style-type: none"> <li>- Feature keysのみ</li> </ul>	<p>配列のアノテーション：</p> <ul style="list-style-type: none"> <li>- Feature keys及びQualifiers</li> </ul>
<p><u>以下の配列の記載が可能</u>：</p> <ul style="list-style-type: none"> <li>- 10個未満の具体的に定義されたヌクレオチド</li> <li>- 4個未満の具体的に定義されたアミノ酸</li> </ul>	<p><u>以下の配列の記載は禁止</u>：</p> <ul style="list-style-type: none"> <li>- 10個未満の具体的に定義されたヌクレオチド</li> <li>- 4個未満の具体的に定義されたアミノ酸</li> </ul>



# WIPO ST.25とST.26の比較

## ...一般情報

ST.25	ST.26
<u>全ての優先権出願情報</u> を記載できる	最先の優先権出願 <u>のみ</u> 記載できる
<u>全ての出願人名及び発明者名</u> を記載できる	出願人1名及び <u>任意</u> で発明者1名 <u>のみ</u> 記載できる
1つの発明の名称が許可されている	それぞれ異なる言語で表示された複数の発明の名称が許可されている
出願人名・発明者名および発明の名称は基本ラテン文字で表記しなければならない	出願人名・発明者名は基本ラテン語の翻訳または翻字を伴う有効なUnicode文字を使用して記載できる

# WIPO ST.25とST.26の比較

## ...配列データ (1)

ST.25	ST.26
DNA、RNAまたはPRTとして特定された配列のみ	分子をさらに詳しく説明するために必須の mol_type qualifierを伴う、DNA、RNAまたはAAとして特定された配列
生物名 : - ラテン語の属名または種名 - ウイルス名 - 「人工配列」 - 「不明」	生物名 : - ラテン語の属名または種名 - ウイルス名 - 「合成構造」 - 「未確認」
「u」は塩基配列中のウラシルを表す	「t」はRNA配列中のウラシルを、DNA配列中のチミンを表す
アミノ酸配列は3文字の略語で表される	アミノ酸配列は1文字の略語で表される

# WIPO ST.25とST.26の比較

## ...配列データ (2)

ST.25	ST.26
「n」および「Xaa」変数はfeatureで提供される定義が必要(must)	定義のない「n」および「X」変数に想定されるデフォルト値
feature locationの形式は明確に定義されていない	feature locationの形式は厳密に定義されている。すべての配列型に「<」と「>」の使用が許可され、「^」、「join」、「order」、「complement」の使用は塩基配列に許可されている
「混合モード」の配列（翻訳されたアミノ酸配列を伴う塩基配列）が認められている	「混合モード」はない。塩基配列の翻訳は「translation」 qualifiersにのみ記載される

# WIPO ST.26への移行

- CWS/5 は**2022年1月1日**を移行日とすることに同意し、その日を「**ビッグバン**」の日と呼んでいる
- **全ての知財庁が国際(PCT)レベル、国内レベル、広域レベルで同時に移行する**
- ST.25 またはST.26のどちらの配列規則に該当するかを決定する基準日は**国際出願日 (IFD)** であり、優先日は基準日とならない。
  - 注：ST.25は2022年1月1日より前の出願に有効である

# WIPO ST.26

## ...記載されるべき事項

- 塩基配列：
  - 10個以上の「具体的に定義された」\* および「列挙された」\* 残基
  - ペプチド核酸（PNA）やグリコール核酸（GNA）などのヌクレオチド類似体配列を含む
- アミノ酸配列：
  - 4個以上の「具体的に定義された」および「列挙された」残基
  - D-アミノ酸を伴う配列を含む
  - 分岐配列の線形部分は配列表に記載する必要がある

\* 定義についてはST.26附属書VIを参照のこと

# WIPO ST.26

## ...記載されるべき事項

- ・ 「具体的に定義された」ヌクレオチド又はアミノ酸とは何か？
- 「具体的に定義された」とは記号「n」で表されるもの以外のヌクレオチド、および記号「X」で表されるもの以外のアミノ酸を意味し、附属書IIに記載されている
- 「特異的に定義された」残基のみが最小長要件を満たす：
  - 10個以上の具体的に定義されたヌクレオチド、または、
  - 4個以上の具体的に定義されたアミノ酸

5'- anctggcaan - 3'

特異的に定義されたヌクレオチドは8個のみ、  
配列表に記載してはならない

5'- agctggcaat - 3'

特異的に定義されたヌクレオチドは10個、  
配列表に記載しなければならない

# WIPO ST.26: XML文書

- 1つのXML 1.0形式のファイルで提供しなければならない
- WIPO ST.26 DTD (附属書II)および標準の内容から派生する事務規定の規則に照らして検証しなければならない
- ユニコードUTF-8を使用してエンコードしなければならない
- ST.26 配列表の構造：
  - XML 宣言  
`<?xml version="1.0" encoding="UTF-8"?>`
  - 文書型 (DOCTYPE) 宣言  
`<!DOCTYPE ST26SequenceListing PUBLIC "-//WIPO//DTD  
Sequence Listing 1.3//EN" "ST26SequenceListing_V1_3.dtd">`
  - ルート要素
    - 一般情報部
    - 配列データ部

# XML の基礎知識

- XML=eXtensible Markup Language（拡張可能なマークアップ言語）
- 情報は記述的要素と属性を使用して「タグ付け」される
- 人間と機械が判読できる標準化したデータ交換手段
- DTD = Document Type Definition（文書型定義） – XML文書の構造及び法的要素や属性を定義する



# XML の基礎知識

## 要素、属性、値 (1)

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

“要素”

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

# XML の基礎知識

## 要素、属性、値 (2)

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

“要素値”

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

# XMLの基礎知識

## 要素、属性、値 (3)

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

“属性”

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

“属性値”

# XML の基礎知識

## 予約文字

要素値に定義済み実体による置換の必要あり

Reserved Character	Predefined Entities
<	&lt;
>	&gt;
&	&amp;
“	&quot;
'	&apos;

例：好ましいfeature location は「<50..62」

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



`<INSDFeature_location>&lt;50..62</INSDFeature_location>`



# WIPO ST.26 : 实例

```

<?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>atgaaattaaaacataaaaarggatgataaaatgagatttgatataaaaagg</INSDSeq_sequence>
    </INSDSeq>
  </SequenceData>
</ST26SequenceListing>

```

# WIPO ST.26 : 構成要素 (1)

1行目 - XML宣言  
(ST.26 第39項(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">
```

# WIPO ST.26 : 構成要素 (2)

2行目 - 文書型 (DOCTYPE) 宣言  
(ST.26 第39項(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">
```

# WIPO ST.26 : 構成要素 (3)

3行目 - ルート要素  
(ST.26 第43項)

```
<?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">
```

一般情報 :

(ST.26 第38項(a),  
第45項から第49項)



# WIPO ST.26 : 構成要素 (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_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>atgaaattaaaacataaaaarggatgataaaatgagatttgatataaaaaagg</INSDSeq_sequence>
  </INSDSeq>
</SequenceData>
</ST26SequenceListing>

```

配列データ  
(ST.26 第38項(b),  
第50項から第100項)

# WIPO ST.26 : 一般情報 (1)

- 「出願の識別」 の項目
  - 出願番号、出願日、知財庁コードが判明していれば必須
  - 不明な場合は出願人が指定する識別番号 (ApplicantFileReference) のみで足りる

```
<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>
```

# WIPO ST.26 : 一般情報 (2)

- 「優先権出願」の項目
  - 配列表に記載可能な優先権出願は1つのみであり、最先の優先権出願でなければならない
  - 優先権が主張されている場合は必須

```
<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>
```

## WIPO ST.26 : 一般情報 (3)

- 「出願人名及び発明者名」の項目
  - 出願人名及び発明者名は、それぞれ1名ずつのみ配列表に記載することができ、それは「筆頭」出願人及び発明者でなければならない
  - 出願人名は必須であるが、発明者名は任意
  - 出願人名及び発明者名の言語コードは必須
  - 出願人名および／または発明者名にUnicodeの基本ラテン文字以外が含まれている場合、翻字または翻訳された基本ラテン文字を記載する必要がある

```
<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>
```

# WIPO ST.26 : 一般情報 (4)

- 「発明の名称」の項目
  - 出願言語で少なくとも1つの発明の名称が必須
  - 他の言語による名称を追加することも可能
  - 各名称には言語コードが必須

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

# WIPO ST.26 : 一般情報 (5)

- 「配列の総数」の要素
  - 必須
  - 総数にはスキップされた配列を含まなければならない

<SequenceTotalQuantity>**6**</SequenceTotalQuantity>

# WIPO ST.26 : 配列 (1)

- 必須要素

「配列識別番号」または「SEQ ID NO」

```

<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>atgaaattaaaacataaaarggatgataaatgagatttgatataaaaaagg</INSDSeq_sequence>
  </INSDSeq>
</SequenceData>

```

配列の長さ

分子型 (DNA, RNA, AA)

区分 (常に「PAT」)

# WIPO ST.26 : 配列 (2)

- 必須の「source」または「SOURCE」 feature

```

<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>atgaaattaaaacataaaaarggatgataaaatgagatttgatataaaaaagg</INSDSeq_sequence>
  </INSDSeq>
</SequenceData>

```

1つの配列につき、1つの  
「source」 featureのみ必要

「source」の存在位置は  
配列全体にわたる

2つの必須Qualifier :  
「mol\_type」及び  
「organism」



# WIPO ST.26 : 配列 (3)

- ST.26 配列データ :

```

<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>

```

# WIPO ST.26 : 配列 (4)

- 必須の「mol\_type」および「MOL\_TYPE」 qualifier – 値の選択肢:

## DNA

ゲノムDNA

その他 DNA

未割り当てDNA

## RNA

ゲノムRNA

mRNA

tRNA

rRNA

その他RNA

転写RNA

ウイルスcRNA

未割り当てRNA

## AA

蛋白質

# WIPO ST.26 : 配列 (5)

- 必須の「organism」および「ORGANISM」 qualifier – 値の選択肢:
  - ラテン語の属名と種名 例 : 「Mus musculus」
  - 属名の後に「sp.」が付くもの 例 : 「Mus sp.」
  - ウイルス名 例 : 「Torque teno virus 1」
  - 「未確認」
  - 「合成構造」
- 「マウス」などの一般名を生物名として使用してはならない。必要に応じて、一般名を配列表のnote qualifierに含めることができる

# WIPO ST.26 : 配列 (6)

## Feature keys および Qualifiers

必須の「source」または「SOURCE」featureに加え、出願人は配列をさらに詳しく説明するために複数の任意のfeaturesを追加できる

塩基配列およびアミノ酸配列では、それぞれ異なるfeature keysを使用する

- 各featureは、1つまたは複数の任意のqualifiersを有する場合もあり、また、1つの必須のqualifierを有する場合もある

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

(WIPO ST.26, 附属書 I, 第 5 節から第 8 節)

# WIPO ST.26 : 配列 (7)

- 塩基配列 :
  - 全て小文字の記号
  - スペースなし、番号なし
  - 「u」記号なし、「t」はRNA中のウラシルを表す
  - 「n」のデフォルト値は「『a』、『c』、『g』、又は『t/u』のいずれか1つ」

Symbol	Nucleotide
a	adenine
c	cytosine
g	guanine
t	thymine in DNA/uracil in RNA (t/u)
m	a or c
r	a or g
w	a or t/u
s	c or g
y	c or t/u
k	g or t/u
v	a or c or g; not t/u
h	a or c or t/u; not g
d	a or g or t/u; not c
b	c or g or t/u; not a
n	a or c or g or t/u; "unknown" or "other"

# WIPO ST.26 : 配列 (8)

- アミノ酸配列 :
  - 全て1文字、大文字の記号
  - スペースなし、番号なし
  - 「X」のデフォルト値は「『A』、『R』、『N』、『D』、『C』、『Q』、『E』、『G』、『H』、『I』、『L』、『K』、『M』、『F』、『P』、『O』、『S』、『U』、『T』、『W』、『Y』又は『V』のいずれか1つ」

Symbol	Amino acid
A	Alanine
R	Arginine
N	Asparagine
D	Aspartic acid (Aspartate)
C	Cysteine
Q	Glutamine
E	Glutamic acid (Glutamate)
G	Glycine
H	Histidine
I	Isoleucine
L	Leucine
K	Lysine
M	Methionine
F	Phenylalanine
P	Proline
O	Pyrrolysine
S	Serine
U	Selenocysteine
T	Threonine
W	Tryptophan
Y	Tyrosine
V	Valine
B	Aspartic acid or Asparagine
Z	Glutamine or Glutamic acid
J	Leucine or Isoleucine
X	A or R or N or D or C or Q or E or G or H or I or L or K or M or F or P or O or S or U or T or W or Y or V; "unknown" or "other"

# WIPO ST.26 : 配列 (9)

- スキップされた配列 : 出願人は後続の配列に番号を付け直すことなく、配列表から配列データを削除できる
  - INSDSeq\_length, INSDSeq\_moltype, INSDSeq\_division は存在するが、値はなし
  - feature tableなし、source featureなし;
  - 配列の要素の値は「000」である必要あり

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

# WIPO ST.26 : 内容

- 本文 – 記載／表現の要件
- 附属書I – INSDCに基づく統制語彙
- 附属書II – ST.26 文書型 DTD
- 附属書III – ST.26配列表XMLファイルの実例
- 附属書IV – ST.26 XML インスタンスに用いる基本ラテン文字コードの文字サブセット
- 附属書V – INDSデータ交換要件 (知財庁のみ)
- 附属書VI – 実例付ガイダンス文書
- 附属書VIの付録 – 附属書VIで例示された全ての配列開示を含む XML ファイル
- 附属書VII – ST.25からST.26への配列表の移行に関する推奨事項



# WIPO ST.26:

## 本文

参照項目	内容
第1項～第9項	はじめに、定義、範囲、参考
第10項～第37項	配列の表現
第38項～第49項	XMLによる配列表の構造
第50項～第71項	配列データ部、Feature table、Feature keys、 必須の Feature keys、Feature location
第72項～第84項	Feature qualifiers、必須のFeature qualifiers
第85項～第100項	フリーテキスト、コード配列、変異体

# WIPO ST.26 附属書I :

## 統制語彙

セクション	内容
1	ヌクレオチドのリスト（小文字、1文字記号）
2	修飾ヌクレオチドのリスト
3	アミノ酸のリスト（大文字、1文字記号）
4	修飾アミノ酸のリスト
5	塩基配列のFeature keys
6	塩基配列のQualifiers
7	アミノ酸のFeature keys（UniProtより引用）
8	アミノ酸配列のQualifiers
9	遺伝コード表

# WIPO ST.26 附属書 II :

## ST.26 DTD

- ST.26 DTDの詳細 : 最新版 1.3
  
- 一般情報部
  - 特許出願情報に関する要素
  
- 配列データ部
  - INSDC DTDのサブセット
  - 各要素に1つの配列に関する情報を含む1つ以上の配列データの要素

# WIPO ST.26 附属書 VI :

## ガイダンス文書

- 配列の開示に関する49の実例と、ST.26規則が各例にどのように適用されるかを説明している
  
- 各例は以下の対応を示す：
  1. 配列が配列表に含まれる必要があるか、含まれることが許可され又は禁止されているか。
  2. 配列が配列表に含まれる必要があるか、または含むことが許可されている場合、それをどのように表現しなければならないか
  
- 附属書VIの付録はガイダンス文書に記載されている全ての実例を含む ST.26 XML配列表である

# WIPO ST.26 附属書 VII :

## ST.25からST.26への配列表の変換に関する推奨事項

- 新規事項が含まれないようにする方法
- ST.25 とは異なるST.26の要件 – ST.26規則では ST.25規則では求められなかった情報が必要になる
- ST.25形式からST.26形式への配列表の変換には、常に出願人からの入力が必要となる
- 附属書VII の推奨事項に従えば、ST.25に準拠した配列表をST.26形式に変換しても新規事項は発生しない
- 推奨事項および実例と共に20の変換シナリオを提示

日本語的に違和感があるところ、推進室にて確認ください。

# WIPO Sequence (1)

- ST.26 に準拠した配列表をオーサリングし、検証し、作成するために、WIPOが開発したデスクトップツール
- WIPOが加盟国からの要請を受け、国際レベル、国内レベル、広域レベルの全ての関連官庁および出願人のために開発した共通ツール
- WIPO Sequenceを使用することで、ユーザーフレンドリーなインターフェイスでST.26 XMLの作成が簡素化され、XMLファイルを直接編集する必要はない
- 最新版の無料ダウンロード：  
<https://www.wipo.int/standards/en/sequence/index.html>

# WIPO Sequence (2)

- 配列情報をプロジェクトに保存し、検証した上で、ST.26形式の配列表を作成することができる
- データはST.26配列表、ST.26プロジェクト、ST.25配列表、multi-sequence formatファイル、raw formatファイル、FASTA formatファイルからインポートできる
- XML形式の配列表の検証もできる
- 関連するfeature keys、qualifiers、生物名はドロップダウンメニューから簡単に選択できる
- 出願人および発明者の情報は「Persons and Organizations」データベースで保存できる
- 翻訳プログラムにより使用される XLIFF ファイルのエクスポートおよびインポートに対応

# WIPO Sequence : プロジェクトトップページ

WIPO | Sequence PROJECTS PERSONS & ORGANIZATIONS ORGANISMS HELP PREFERENCES ENGLISH

NEW PROJECT IMPORT PROJECT IMPORT SEQUENCE LISTING VALIDATE SEQUENCE LISTING

## PROJECTS


Search project by name

Project name :	Applicant file reference	Applicant name	Invention title	Status	Creation Date :
<a href="#">16079428_no_source_aa_gid1</a>	H0075.70243US00	Ferring B.V.	STABLE LIQUID GONADOTROPIN FORMULATION	invalid	2021-02-01
<a href="#">All features and qualifiers</a>	123abc	Simple Healthkit, Inc.	Compositions and Methods for Treating Cancer	modified	2021-02-01
<a href="#">Ark three letter aa code import from ST.25</a>	abc123	Joe, Smith	Improper 3-letter AA code for import	new	2021-02-26
<a href="#">Beta 4 Test Project</a>	abc123	Stephenson, Adam	Beta 4 Test Project	modified	2021-03-11
<a href="#">DNA RNA with u and t</a>	uspto	Biotech, Inc.	Beta 4 test import ST25	modified	2021-03-15



# WIPO Sequence : プロジェクトの詳細

## 一般情報の項目


 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)	<a href="#">PCT/IB2014/111111</a>	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	
--	-----------------------	---

# WIPO Sequence : プロジェクトの詳細

## 配列の項目

WIPO | Sequence
Return to project home

FOR WIPO TRAINING
VERIFICATION REPORT
FREE TEXT QUALIFIERS
IMPORT REPORT
DISPLAY THE SEQUENCE LISTING
HELP
PREFERENCES
ENGLI

GENERAL INFORMATION SEQUENCES

---

### SEQUENCE 1

Sequence Number (ID) 1	Molecule Type DNA
Sequence Name test	Organism Mus musculus
Length 52	

▼ FEATURES

Add feature

Feature Key	Location	Qualifiers
<a href="#">source</a>	1..52	mol_type = genomic DNA organism = Mus musculus

▼ SEQUENCE

```
atgaaattaa aacataaaaar ggatgataaa atgagatttg atataaaaaa gg
```

< 1/1 >

# 次回のウェビナー

[standards@wipo.int](mailto:standards@wipo.int)

# 質疑応答

# 用語集:頭字語

- CWS: Committee on WIPO Standards (WIPO標準委員会)
- DDBJ: DNA Databank of Japan (日本DNAデータバンク)
- EMBL-EBI: The European Bioinformatics Institute (欧州バイオインフォマティクス研究所)
- EPO: European Patent Office (欧州特許庁)
- INSDC: International Nucleotide Sequence Database Collaboration (国際塩基配列データベース連携)
- IPO: Intellectual Property Office (知的財産庁)
- NCBI: National Center for Biotechnology Information (アメリカ国立生物工学情報センター)
- WIPO: World Intellectual Property Organization (世界知的所有権機関)