

# ■ WIPO 표준 ST.26 소개

웨비나 교육

# 오늘 다룰 주제

- 서열 목록이란?
- INSDC 서열 데이터베이스
- 새로운 표준이 필요한 이유는?
- WIPO ST.26의 이점
- ST.25 대 ST.26 – 차이점은?
- ST.26으로 전환
- XML 기초
- ST.26 기초– 서열 목록의 일부
- WIPO ST.26의 내용
- WIPO 서열 소개

# 서열 목록이란?

서열 목록은...

- 특허 출원에 개시된 뉴클레오티드 및 / 또는 아미노산 서열을 포함하고 발명의 설명의 일부를 구성합니다.
- 주석으로 알려진 각 서열에 대한 설명 정보를 포함합니다.
- 관련 WIPO 표준의 요구 사항을 준수합니다.(ST.25 or ST.26)
- 발명의 서열 데이터를 검색할 수 있도록 합니다:
  - 특허청 내부
  - 공개적으로 사용 가능한 데이터베이스(INSDC 데이터베이스)

# 배경- INSDC

- INSDC: 국제 뉴클레오티드 서열 데이터베이스 협업:
  - DDBJ: 일본 DNA 데이터뱅크
  - EMBL-EBI: 유럽 생물정보학 연구소
  - NCBI: 국립 생명공학 정보센터(GenBank)
  
- INSDC 데이터베이스에 공개/등록된 출원 서열 데이터를 제출하는 특허청에는 다음이 포함됩니다:
  - 유럽 특허청
  - 일본 특허청
  - 한국 특허청
  - 미국 특허상표청
  
- INSDC 데이터베이스는 공개적으로 검색 가능합니다.

# 새로운 표준이 필요한 이유?

- 현재 서열 목록은 WIPO ST.25에 따라 제출됩니다.

하지만...

- ST.25 포맷은 INSDC 요구사항을 준수하지 않으므로 공용 데이터베이스에 입력하면 데이터가 손실됩니다.
- ST.25 규칙은 명확하지 않으며 전 세계 특허청들은 규칙을 다르게 해석하고 시행합니다.
- 오늘날 일반적인 서열 유형은 ST.25 규칙(뉴클레오타이드 유사체, D- 아미노산, 분지 서열)에 포함되지 않으므로 검색 가능한 데이터베이스에 존재하지 않습니다.
- 데이터가 구조화되지 않음 – ST.25 포맷은 자동 검증 및 데이터 교환에 사용하기 어렵습니다.

# WIPO ST.26의 이점(1)

- 전 세계 특허청에서 표준화된 공통된 서열 목록 포맷을 사용\*
- 표준은 서열 규칙 적용에 대한 특허청 간의 합의를 보장하는 지침 역할을 합니다.
- 서열 목록에 포함되도록 요구되거나 허용되는 서열 공개와 이러한 서열을 표현하는 방법을 명확히 합니다.
- XML 서열 목록의 구조로 인해 제출 품질 향상
- 데이터 검증 자동화 향상 및 특허청의 간소화된 처리

\* 특정 특허청의 경우, 언어 종속 자유 텍스트 한정자를 출원 언어로 번역해야 하는 경우는 예외이며, 대체 서열 목록이 필요할 수 있습니다.

# WIPO ST.26의 이점(2)

- INSDC 데이터베이스 공급자 요구사항과의 데이터 호환성-DDBJ, EBI 및 NCBI – 서열 주석(기능 키 및 한정자)이 공개적으로 검색 가능한 데이터베이스에 포함됩니다.
  
- 표준화:
  - 기능 주석
  - 기능 위치
  - 한정자 및 한정자 값
  - 서열 변형 프레젠테이션
  
- 추가 유형의 서열(뉴클레오타이드 유사체, D- 아미노산, 분지 서열)을 포함해야 하는 요구사항은 더 많은 서열 데이터를 검색할 수 있음을 의미합니다.

# WIPO ST.25 대 ST.26

ST.25	ST.26
숫자 식별자를 갖는 ASCII .txt	구성요소와 속성을 갖는 XML
<p>다음을 포함할 <b>필요가 없음</b>:</p> <ul style="list-style-type: none"> <li>- D- 아미노산</li> <li>- 분기 시퀀스의 선형 부분</li> <li>- 뉴클레오티드 유사체</li> </ul>	<p>다음을 <b>포함해야 함</b>:</p> <ul style="list-style-type: none"> <li>- D- 아미노산</li> <li>- 분기 시퀀스의 선형 부분</li> <li>- 뉴클레오티드 유사체</li> </ul>
<p>서열 주석:</p> <ul style="list-style-type: none"> <li>- 기능 키만</li> </ul>	<p>서열 주석:</p> <ul style="list-style-type: none"> <li>- 기능 키 및 한정자</li> </ul>
<p>서열 포함 <u>허용</u>:</p> <ul style="list-style-type: none"> <li>- &lt; 10 개의 특별히 정의된 뉴클레오티드</li> <li>- &lt; 4개의 특별히 정의된 아미노산</li> </ul>	<p><u>금지된</u> 서열:</p> <ul style="list-style-type: none"> <li>- &lt; 10개의 특별히 정의된 뉴클레오티드</li> <li>- &lt; 4개의 특별히 정의된 아미노산</li> </ul>

# WIPO ST.25 대 ST.26

## ...일반 정보

ST.25	ST.26
모든 우선권 출원 정보가 포함될 수 있습니다.	가장 빠른 우선권 출원만 포함될 수 있습니다.
모든 출원인 및 발명자 명칭이 포함될 수 있습니다.	한 명의 출원인 및 선택적으로 한 명의 발명자만 포함될 수 있습니다.
하나의 발명의 명칭이 허용	각기 다른 언어로 된 여러 개의 발명의 명칭이 허용됨
출원인/발명자 명칭 및 발명의 명칭은 기본 라틴 문자여야 합니다.	출원인/발명자 명칭은 기본 라틴어 번역 또는 음역과 함께 유효한 유니코드 문자를 사용하여 포함될 수 있습니다.

# WIPO ST.25 대 ST.26

## ...서열 데이터(1)

ST.25	ST.26
DNA, RNA 또는 PRT로만 확인된 서열	분자를 추가로 설명하기 위해 필수 mol_type 한정자와 함께 DNA, RNA 또는 AA로 식별된 서열
유기체 명칭: <ul style="list-style-type: none"> <li>- 라틴 속/종</li> <li>- 바이러스 명칭</li> <li>- “인공 서열”</li> <li>- “미확인”</li> </ul>	유기체 명칭: <ul style="list-style-type: none"> <li>- 라틴 속/종</li> <li>- 바이러스 명칭</li> <li>- “합성 구조”</li> <li>- “미확인”</li> </ul>
"u"는 뉴클레오티드 서열에서 우라실을 나타냅니다.	"t"는 RNA 서열에서 우라실을 나타내고 DNA 서열에서 티민을 나타냅니다.
세 글자 약어로 표시되는 아미노산 서열	한 문자 약어로 표시되는 아미노산 서열

# WIPO ST.25 대 ST.26

## ...서열 데이터(2)

ST.25	ST.26
"n" 및 "Xaa" 변수에는 기능에 제공된 정의가 있어야 합니다.	정의가 없는 "n" 및 "X" 변수에 대해 가정되는 기본값
명확하게 정의되지 않은 기능 위치 형식	엄격하게 정의된 기능 위치 형식; 모든 서열 유형에서 "<and>" 를 사용하고 뉴클레오티드 서열에서 "^", "join", "order" 및 "complement"를 사용할 수 있음
허용되는 "혼합 모드" 서열 – 아래 표시된 아미노산 번역이 있는 뉴클레오티드 서열	혼합 모드 없음; 뉴클레오티드 번역은 "번역" 한정자에만 포함됩니다.

# WIPO 표준 ST.26으로 전환

- WIPO표준위원회 제5차 회의에서 회원국은 ‘빅뱅’일로 불리는 **2022년 1월 1일**의 전환일에 합의
- 전환은 모든 지식재산청(IPO)이 국제(PCT출원), 국가 및 지역 수준에서 동시에 전환됩니다.
- 출원일 (우선권일)이 전환의 기준일. 즉 출원일에 따라 ST.25 또는 ST.26 서열 규칙에 해당하는지 여부를 결정
- 참고 : ST.25는 출원일이 2022년 1월 1일 이전인 출원에 대해 유효합니다.

# WIPO ST.26

...무엇이 반드시 포함되어야 할까요?

- 뉴클레오타이드 서열:
  - 10개 이상의 “구체적으로 정의” \* 되고 “열거된”\* 잔기
  - 펩티드 핵산(PNA) 및 글리콜 핵산(GNA)과 같은 뉴클레오타이드 유사체가 있는 서열 포함
- 아미노산 서열:
  - 4개 이상의 “구체적으로 정의”되고 “열거된”\* 잔기
  - D-아미노산 서열 포함
  - 분기된 서열의 선형 영역이 서열 목록에 포함되어야 함

*\*관련 용어 정의는 ST.26 부록 VI 참조*

# WIPO ST.26

## ...무엇이 반드시 포함되어야 할까요?

- “특별하게 정의”된 뉴클레오티드 또는 아미노산은 무엇일까요?
- “특별하게 정의 ”란 의미는 부록 I에 나열된 기호 "n"으로 표시되는 것 이외의 모든 뉴클레오티드 및 기호 "X"로 표시되는 것 이외의 아미노산을 의미합니다.
- “특별하게 정의된” 잔기만 최소 길이 요건에 포함됩니다:
  - 10개 이상의 뉴클레오티드; 또는,
  - 4개 이상의 특별하게 정의된 아미노산

5'- anctggcaan – 3'

단지 8개의 특별하게 정의된 뉴클레오티드 존재; 서열 목록에 포함되지 않아야 합니다.

5'- agctggcaat – 3'

10개의 특별하게 정의된 뉴클레오티드 존재; 서열 목록에 포함되어야 합니다.

# WIPO ST.26: XML 문서

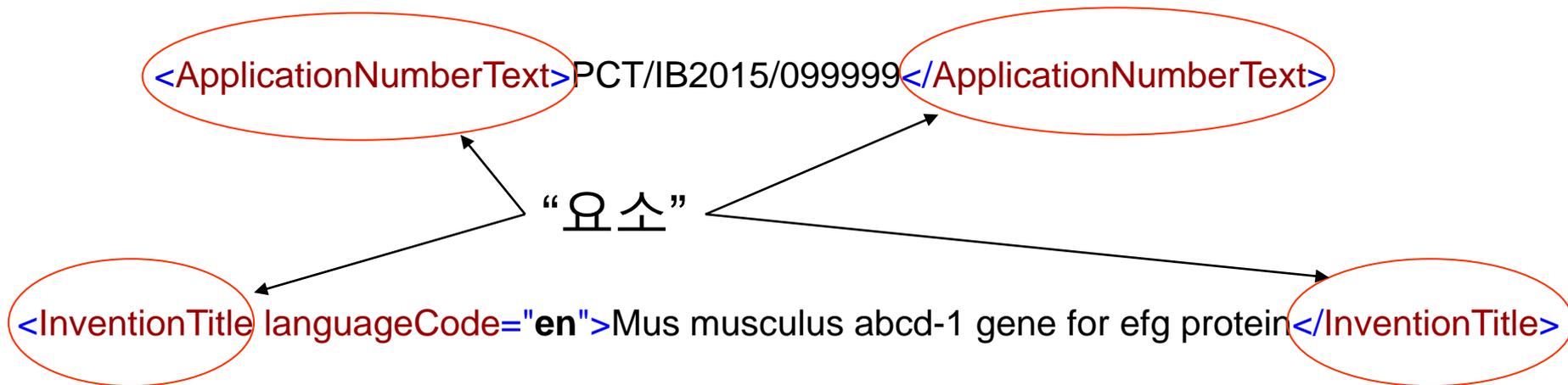
- 하나의 XML 1.0 형식 파일로 제공되어야 합니다.
- WIPO ST.26 DTD(부록 II) 및 표준 내용에서 파생된 비즈니스 규칙에 대해 검증해야 합니다.
- 유니 코드 UTF-8을 사용하여 인코딩해야 합니다.
- 서열 목록의 구조:
  - XML선언
  - 문서 유형(DOCTYPE) 선언
  - 루트 요소
    - 일반 정보 부분
    - 서열 데이터 부분

# XML DTD 기초

- XML = eXtensible Markup Language
- 정보는 설명 요소와 속성을 사용하여 '태그'됩니다.
- 사람과 기계가 모두 읽을 수 있는 표준화된 데이터 교환 수단
- DTD = Document Type Definition - XML 문서의 구조와 법적 요소 및 속성을 정의합니다.

# XML 기초

## 요소, 속성 및 값(1)



# XML 기초

## 요소, 속성 및 값(2)

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

“요소 값”

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

# XML 기초

## 요소, 속성 및 값(3)

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

“속성”

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

“속성 값”

# XML 기초

## 예약된 문자

요소 값에서 사전 정의된 엔티티로 대체되어야 함

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

예: 원하는 기능 위치는 “<50..62”

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



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



# WIPO ST.26: 예

```

<?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: 구성요소(1)

1행 - XML 선언  
(ST.26 문단 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">
```

# WIPO ST.26: 구성요소(2)

2행 - 문서 유형(DOCTYPE) 선언  
(ST.26 문단 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">
```

# WIPO ST.26: 구성요소(3)

3행 - 루트 요소  
(ST.26 문단 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">
```

일반 정보:

(ST.26 문단. 38(a),  
45-49)

# WIPO ST.26: 구성요소(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>atgaaattaaacataaaaarggatgataaaatgagatttgatataaaaaagg</INSDSeq_sequence>
  </INSDSeq>
</SequenceData>
</ST26SequenceListing>

```

서열 데이터

(ST.26 문단. 38(b),  
50-100)

# WIPO ST.26: 일반 정보(1)

- “출원인 식별” 섹션
  - 출원번호, 출원일 및 지식재산청 코드는 알려진 경우 필수;
  - 그렇지 않으면 출원인 파일 참조만으로 충분합니다.

```

<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: 일반 정보(2)

- “우선권 출원” 섹션
  - 서열 목록에는 단 하나의 우선권 출원만 포함될 수 있으며, 가장 빠른 우선권 출원이어야 합니다;
  - 우선권이 주장되는 경우, 필수

```

<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: 일반 정보(3)

- “출원인 및 발명자 명칭” 섹션
  - 하나의 출원인 명칭과 하나의 발명자 명칭만 서열 목록에 포함될 수 있으며 반드시 “주요” 출원인 및 발명자이어야 합니다;
  - 출원인 명칭은 필수이고; 발명자 명칭은 선택사항입니다;
  - 출원인 및 발명자 명칭의 언어 코드는 필수입니다;
  - 출원인 및/또는 발명자 명칭에 비 유니코드 기본 라틴문자가 포함된 경우, 기본 라틴 문자로의 음역 또는 번역이 포함되어야 합니다.

```

<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: 일반 정보(4)

- “발명의 제목” 섹션
  - 출원 언어로 된 하나 이상의 발명의 제목은 필수입니다;
  - 다른 언어로 된 추가 제목이 포함될 수 있습니다;
  - 각 제목에는 언어 코드가 필수입니다;

```
<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: 일반 정보(5)

- “서열 총 수량” 요소
  - 필수;
  - 총계는 건너 뛴 서열을 포함해야 합니다.

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

# WIPO ST.26: 서열(1)

- 필수 요소

```

<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>

```

“서열 식별 번호”  
또는 “SEQ ID NO:”

서열 길이

분자 유형(DNA, RNA, or AA)

분할(항상 “PAT”)

# WIPO ST.26: 서열(2)

- 필수 “소스” 또는 “소스” 기능

```

<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>

```

서열당 하나의 “소스” 기능만 필요

“소스” 위치는 전체 서열에 걸쳐  
있어야 합니다.

두 가지 필수 한정자:  
“mol\_type” 및  
“유기체”

# WIPO ST.26: 서열(3)

ST.26 서열 데이터:

```
<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>atgaaattaaaacataaaarggatgataaaatgagatttgatataaaaaagg</INSDSeq_sequence>
  </INSDSeq>
</SequenceData>
```

# WIPO ST.26: 서열(4)

- 필수 “mol\_type” 및 “MOL\_TYPE” 한정자 – 값 선택:

## DNA

게놈 DNA  
기타 DNA  
미할당 DNA

## RNA

게놈 RNA  
mRNA  
tRNA  
rRNA  
기타 RNA  
전사된 RNA  
바이러스 cRNA  
미할당 RNA

## AA

단백질

# WIPO ST.26: 서열(5)

- 필수 “유기체” 및 “ORGANISM” 한정자- 값 선택:
  - 라틴 속 및 종 이름(예: “Mus musculus”)
  - 속 이름 뒤에 “sp”가 붙습니다.(예: “Mus sp”)
  - 바이러스 이름(예: “Torque teno virus 1”)
  - “미확인”
  - “합성 구조”
- “마우스”와 같은 일반적인 이름은 유기체 이름으로 사용해서는 안됩니다. 원하는 경우, 노트 한정자의 서열 목록에 일반 이름을 포함할 수 있습니다.

# WIPO ST.26: 서열(6)

## 기능 키 및 한정자

필수 “소스” 또는 “SOURCE“ 기능 외에도, 출원인은 목록을 상세히 설명하기 위해 여러 선택적 기능을 추가할 수 있습니다.

- 뉴클레오타이드 서열 및 아미노산 서열에 대한 상이한 기능 키;
- 각 기능에는 하나 이상의 선택적 한정자가 있을 수 있으며, 필수 한정자가 있을 수 있습니다.

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, Annex I, Sections 5-8)

# WIPO ST.26: 서열(7)

- 뉴클레오타이드 서열:
  - 모든 소문자 기호;
  - 공백 없음, 번호 매기기 없음;
  - “u” 기호 없음; “t”는 RNA에서 우라실을 나타냄
  - “n”의 기본값은 “a’, ‘c’, ‘g’, 또는 ‘t/u’ 중 하나” 입니다.

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: 서열(8)

- 아미노산 서열:
  - 모든 단일 문자, 대문자, 기호;
  - 공백 없음, 번호 매기기 없음;
  - “X”의 기본값은 ‘A’, ‘R’, ‘N’, ‘D’, ‘C’, ‘Q’, ‘E’, ‘G’, ‘H’, ‘I’, ‘L’, ‘K’, ‘M’, ‘F’, ‘P’, ‘O’, ‘S’, ‘U’, ‘T’, ‘W’, ‘Y’, 또는 ‘V’ 중 하나입니다.

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: 서열(8)

- 건너 뛴 서열: 출원인이 하위 서열의 번호를 다시 매길 필요 없이 서열 목록에서 서열 데이터를 삭제할 수 있습니다.
  - INSDSeq\_length, INSDSeq\_moltype, INSDSeq\_division 이 있지만 값이 없음;
  - 기능 테이블과 소스 기능이 없음;
  - 서열 요소의 값은 “000” 이어야 함

```

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

```

# WIPO ST.26: 목차

- **본문** – 포함/표시에 대한 요건
- **부록 I** – INSDC에 기반한 통제된 어휘
- **부록 II** – ST.26 문서 유형 DTD
- **부록 III** – 사례 ST.26 서열 목록 XML 파일
- **부록 IV** – ST.26 XML 인스턴스에 대한 기본 라틴 코드의 문자 하위 집합
- **부록 V** – INDS 데이터 교환 요건(지식재산청만 해당)
- **부록 VI** – 사례가 포함된 지침 문서
- **부록 VI에 대한 부속서** – 부록 VI에 예시된 모든 서열 공개를 포함하는 XML 파일
- **부록 VII** – ST.25에서 ST.26으로 서열 목록 변환에 대한 권장 사항

# WIPO ST.26:

## 본문

문단 참조	내용
1-9	소개; 정의; 범위; 참고문헌
10-37	목록의 표현
38-49	XML의 서열 목록 구조
50-71	서열 데이터 부분; 기능 테이블; 기능 키; 필수 기능 키; 기능 위치
72-84	기능 한정자; 필수 기능 한정자
85-100	자유 텍스트; 코딩 서열; 변형

# WIPO ST.26 부록 I:

## 통제된 어휘

섹션	내용
1	뉴클레오티드 목록(소문자, 단일 문자 기호)
2	변형된 뉴클레오티드 목록
3	아미노산 목록(대문자, 단일 문자 기호)
4	변형된 아미노산 목록
5	뉴클레오티드 서열의 기능 키
6	뉴클레오티드 서열에 대한 한정자
7	아미노산의 기능 키(UniProt에서 채택)
8	아미노산 서열에 대한 한정자
9	유전 코드 테이블

# WIPO ST.26 부록 II:

## ST.26 DTD

- ST.26 DTD의 세부사항: 현재 버전 1.3
- 일반 정보 부분
  - 특히 출원 정보와 관련된 요소
- 서열 데이터 부분
  - INSDC DTD의 하위 집합
  - 각 요소가 하나의 서열에 대한 정보를 포함하는 하나 이상의 서열 데이터 요소

# WIPO ST.26 부록 VI: 사용안내 문서

- 서열 공개의 실제 사례 49개와 ST.26 규칙이 각 사례에 어떻게 적용되는지에 대한 설명이 포함되어 있습니다;
- 각 사례는 다음을 다룹니다:
  1. 서열이 요구되는지, 허용되는지 또는 서열 목록에 서열의 포함이 금지되는지 여부;
  2. 서열이 서열 목록에 포함되어야 하거나 포함되도록 허용된 경우, 서열의 표시 방법
- 부록 VI의 부속서는 지침 문서에 제시된 모든 예를 포함하는 ST.26 XML 서열 목록입니다.

# WIPO ST.26 부록 VII:

## ST.25 서열 목록을 ST.26으로 변환 시 권장사항

- ...또는 새로운 물질의 포함을 피하는 방법
- ST.26의 요구사항은 ST.25와 다릅니다 – ST.26 규칙은 ST.25 규칙에서 요구하지 않는 정보를 요구합니다.
- 서열 목록을 ST.25 형식에서 ST.26 형식으로 변환하려면 항상 출원인의 입력이 필요합니다.
- 호환 ST.25 서열 목록을 ST.26 형식으로 변환해도 부록 VII의 권장사항을 따르는 경우, 새로운 문제가 발생하지 않습니다.
- 권장사항 및 사례와 함께 제공되는 20개의 변환 시나리오

# WIPO 서열(1)

- ST.26 호환 서열 목록의 작성, 유효성 검증 및 생성을 지원하기 위해 WIPO에서 개발한 데스크탑 도구
- 회원국들은 국제, 국가 및 지역 레벨에서 모든 특허청 및 출원인을 위한 공통 도구의 개발을 WIPO에 요청했습니다.
- WIPO 서열을 사용하면 사용자 친화적인 인터페이스로 ST.26 XML 생성이 단순화됩니다. XML 파일을 직접 편집할 필요가 없습니다.
- 최신 버전을 무료로 다운로드 하세요:

<https://www.wipo.int/standards/en/sequence/index.html>

# WIPO 서열(2)

- 서열 정보를 프로젝트에 저장하고 유효성을 검증한 다음 ST.26 형식의 서열 목록을 생성할 수 있습니다.
- 데이터는 ST.26 서열 목록, ST.26 프로젝트, ST.25 서열 목록, 다중 서열 형식 파일, 원시 형식 파일 및 FASTA 형식 파일에서 가져올 수 있습니다.
- XML 형식의 서열 목록 검증
- 드롭 다운 메뉴에서 관련 기능 키, 한정자 및 유기체 이름을 쉽게 선택할 수 있습니다.
- 출원인 및 발명자 정보는 “개인 및 기관” 데이터베이스에 저장될 수 있습니다.
- 번역자가 사용하는 XLIFF 파일의 내보내기 및 가져오기를 지원합니다

# WIPO 서열: 프로젝트 홈페이지


**WIPO** | Sequence
 PROJECTS
PERSONS & ORGANIZATIONS
ORGANISMS
HELP

 PREFERENCES
 ENGLISH

[NEW PROJECT](#)
[IMPORT PROJECT](#)
[IMPORT SEQUENCE LISTING](#)
[VALIDATE SEQUENCE LISTING](#)

## PROJECTS

Project name ◯	Applicant file reference	Applicant name	Invention title	Status	Creation Date ◯
<a href="#">16079428_no_source_sequence</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 서열: 프로젝트 세부 정보

## 일반 정보 섹션

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 abod-1 gene for efg protein	Language en - English	✎
--	-----------------------	---

# WIPO 서열: 프로젝트 세부 정보

## 서열 섹션

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

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
<u>source</u>	1..52	mol_type = genomic DNA organism = Mus musculus

▼ SEQUENCE

atgaaattaa aacataaaaar ggatgataaa atgagatttg atataaaaaa gg	52	✎
< 1/1 >		

↑

향후 계획은?

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

# Q&A 세션

# 용어집: 약어

- CWS: WIPO 표준위원회
- DDBJ: 일본 DNA 데이터뱅크
- EMBL-EBI: 유럽 생물정보학 연구소
- EPO: 유럽 특허청
- INSDC: 국제 뉴클레오티드 서열 데이터베이스 협업
- IPO: 지식재산청
- NCBI: 국립 생명공학 정보 센터
- WIPO: 세계지식재산기구