

■ WIPO Standard ST.26

上級編

ウェビナートレーニング

注：これは後続モジュールで、参加者はすでにST.26基礎コースに参加していることを前提としています。

ST.26基礎コースはこちらからアクセスできます：

https://www.wipo.int/meetings/en/details.jsp?meeting_id=62848

本日の内容

- よく使用されるfeature keyとqualifier
- featureの位置の形式
- qualifierの値の形式と非英語qualifierの値
- 特別なケース - DNA中のウラシルとRNA中のチミン ; DNA/RNAハイブリッド分子
- ヌクレオチド類似体、D-アミノ酸、および分岐配列
- 変異体配列

よく使用されるfeature key とqualifier

feature key と qualifier

- feature keyを使用することで、位置によって識別される配列の1つ以上の残基についてさらに記述できます
 - ヌクレオチド配列のfeature keyは、Annex IのSection 5に記載されています
 - ヌクレオチド配列のfeature keyは小文字です：例「misc_binding」
 - アミノ酸配列のfeature keyは、Annex IのSection 7に記載されています
 - アミノ酸配列のfeature keyは大文字です：例「REGION」

- qualifierを使用することで、featureについてさらに記述することができます
 - ヌクレオチド配列のqualifierはAnnex IのSection 6に記載されています
 - ヌクレオチド配列のqualifierは小文字です：例「allele」
 - アミノ酸配列のqualifierはAnnex IのSection 8に記載されています
 - アミノ酸配列のqualifierは大文字です：例「NOTE」

feature key & qualifier

```

<SequenceData sequenceIDNumber="17">
- <INSDSeq>
  <INSDSeq_length>7</INSDSeq_length>
  <INSDSeq_moltype>AA</INSDSeq_moltype>
  <INSDSeq_division>PAT</INSDSeq_division>
- <INSDSeq_feature-table>
  - <INSDFeature>
    <INSDFeature_key>SOURCE</INSDFeature_key>
    <INSDFeature_location>1..7</INSDFeature_location>
  - <INSDFeature_quals>
    - <INSDQualifier>
      <INSDQualifier_name>ORGANISM</INSDQualifier_name>
      <INSDQualifier_value>Homo sapiens</INSDQualifier_value>
    </INSDQualifier>
    - <INSDQualifier>
      <INSDQualifier_name>MOL_TYPE</INSDQualifier_name>
      <INSDQualifier_value>protein</INSDQualifier_value>
    </INSDQualifier>
  </INSDFeature_quals>
</INSDFeature>
- <INSDFeature>
  <INSDFeature_key>VARIANT</INSDFeature_key>
  <INSDFeature_location>1</INSDFeature_location>
- <INSDFeature_quals>
  - <INSDQualifier>
    <INSDQualifier_name>NOTE</INSDQualifier_name>
    <INSDQualifier_value>X can be any amino acid</INSDQualifier_value>
  </INSDQualifier>
</INSDFeature_quals>
</INSDFeature>
</INSDSeq_feature-table>
  <INSDSeq_sequence>XYEKGJL</INSDSeq_sequence>
</INSDSeq>
</SequenceData>

```

feature key と qualifier

- feature keyは、すべての配列で必須である「source」 / 「SOURCE」 featureを除き、任意です。
- 各feature keyには、featureについてさらに記述するために使用できるqualifierのリストがあります。ほとんどのqualifierは任意ですが、一部のfeature keyには必須qualifierがあります。

5.31. Feature Key	regulatory
Definition	any region of a sequence that functions in the regulation of transcription, translation, replication or chromatin structure;
Mandatory qualifiers	regulatory_class
Optional qualifiers	allele bound_moiety function gene gene_synonym map note operon phenotype pseudo pseudogene standard_name

- 「source」 / 「SOURCE」 featureには、qualifier 「mol_type」 / 「MOL_TYPE」 および 「organism」 / 「ORGANISM」 が必須です。

feature key と qualifier

- 一部のfeature keyには追加の制約があります
 - 生物の範囲；例えば「C_Region」は真核生物に限られる
 - 分子の範囲；例えば「D-loop」はDNA配列に限られる

5.4. Feature Key	D-loop
Definition	displacement loop; a region within mitochondrial DNA in which a short stretch of RNA is paired with one strand of DNA, displacing the original partner DNA strand in this region; also used to describe the displacement of a region of one strand of duplex DNA by a single stranded invader in the reaction catalyzed by RecA protein
Optional qualifiers	allele gene gene_synonym map note
Molecule scope	DNA

feature key と qualifier

ヌクレオチド配列：「modified_base」 feature key

- 修飾されたヌクレオチドを記述するには、feature key 「modified_base」とその必須 qualifier 「mod_base」を使用する必要があります（ST.26, paragraph 16）
- 「修飾ヌクレオチド」とは、次のものを除くヌクレオチドです：
 - デオキシ-[a, g, c, または t] 3'-ーリン酸
 - [a, g, c, または u] 3'-ーリン酸（ST.26, paragraph 3(f)）
- 「修飾ヌクレオチド」は、可能な限り、対応する未修飾ヌクレオチド（Annex 1, Section 1, Table 1）で表す必要があります。それ以外の場合は「n」で表すことができます。例えば、「2'-O-メチルシチジン」は、配列中、「c」とします。「クエオシン」は「n」とします。記号「n」は、1つの残基のみに対応します。
- 必須の qualifier 「mod_base」の値は、Annex 1, Section 2のTable 2の値から選択する必要があります。その値が「other」の場合、追加の「note」 qualifierは、修飾された残基の省略されていない完全な名前を含んでいなければなりません。

feature keyとqualifier

ヌクレオチド配列：「modified_base」 feature key

- 例：15位にイノシンを含むヌクレオチド配列

```
<INSDFeature>
  <INSDFeature_key>modified_base</INSDFeature_key>
  <INSDFeature_location>15</INSDFeature_location>
  <INSDFeature_qual>
    <INSDQualifier>
      <INSDQualifier_name>mod_base</INSDQualifier_name>
      <INSDQualifier_value>i</INSDQualifier_value>
    </INSDQualifier>
  </INSDFeature_qual>
</INSDFeature>
```

- 「イノシン」は、Annex I, Section 2のTable 2に「i」の略号で記載されています

feature keyとqualifier

ヌクレオチド配列：「modified_base」 feature key

- 例：22位にキサンチンを含むヌクレオチド配列

```
<INSDFeature>
  <INSDFeature_key>modified_base</INSDFeature_key>
  <INSDFeature_location>22</INSDFeature_location>
  <INSDFeature_qual>
    <INSDQualifier>
      <INSDQualifier_name>mod_base</INSDQualifier_name>
      <INSDQualifier_value>OTHER</INSDQualifier_value>
    </INSDQualifier>
    <INSDQualifier>
      <INSDQualifier_name>note</INSDQualifier_name>
      <INSDQualifier_value>xanthine</INSDQualifier_value>
    </INSDQualifier>
  </INSDFeature_qual>
</INSDFeature>
```

- 「キサンチン」は、Annex I, Section 2のTable 2に記載されていません。したがって、必須の「mod_base」 qualifierの値は「OTHER」で、値が「xanthine」の追加の「note」 qualifierを含める必要があります。

feature keyとqualifier

ヌクレオチド配列：「modified_base」 feature key

- 「modified_base」は、脱塩基部位を記述するためにも使用できます：

```
<INSDFeature>
  <INSDFeature_key>modified_base</INSDFeature_key>
  <INSDFeature_location>11</INSDFeature_location>
  <INSDFeature_qual>
    <INSDQualifier>
      <INSDQualifier_name>mod_base</INSDQualifier_name>
      <INSDQualifier_value>OTHER</INSDQualifier_value>
    </INSDQualifier>
    <INSDQualifier>
      <INSDQualifier_name>note</INSDQualifier_name>
      <INSDQualifier_value>abasic site</INSDQualifier_value>
    </INSDQualifier>
  </INSDFeature_qual>
</INSDFeature>
```

- 脱塩基部位は、配列内の「n」で表すことができ、さらに「modified_base」 feature keyと、値が「OTHER」の「mod_base」 qualifierおよび値が「abasic site」の追加の「note」 qualifierとを使用して記述できます。

feature key と qualifier

ヌクレオチド配列：「CDS」 feature key

- feature key 「CDS」を使用することで、コード配列を識別することができます。CDS featureの位置には、終止コドンが含まれていなければなりません。
(ST.26, pgh.89)
- 「CDS」 feature keyには必須のqualifierはありません
- 次のqualifierが一般的に使用されます：

“pseudo”

“pseudogene”

“translation”

“transl_table”

“codon_start”

“transl_except”

“protein_id”

feature key と qualifier

ヌクレオチド配列：「CDS」 feature key

- 「CDS」 featureには、次のqualifierのうちの1つのみを指定できます：

Qualifier	説明	値
pseudo	CDS featureに機能はなく、翻訳もないが、偽遺伝子ではないことを示す	なし
pseudogene	CDS featureは偽遺伝子であり、翻訳がないことを示す	processed unprocessed unitary allelic unknown
translation	CDSの翻訳で得られたアミノ酸配列を示す	1文字のアミノ酸の略号

feature key と qualifier

ヌクレオチド配列：「CDS」 feature key

- コード配列によってコードされ、「translation」 qualifierに示された4個以上の具体的に定義されたアミノ酸を含むアミノ酸配列は、別の配列として配列表に含められ、独自の配列識別番号が割り当てられる必要があります。
- アミノ酸配列に割り当てられた配列識別番号を、「CDS」 feature key内の「protein_id」 qualifierの値として記述する必要があります。
- アミノ酸配列の「SOURCE」 feature keyの「ORGANISM」 qualifierは、そのコード配列のものと同じでなければなりません。（ST.26, pgh.92）

feature keyとqualifier

ヌクレオチド配列：「CDS」 feature key

■ 翻訳された配列を修正できるqualifier：

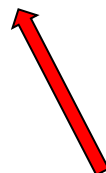
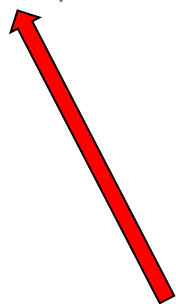
qualifier	説明	値
transl_table	CDSの翻訳に使用される遺伝暗号表を示す；デフォルトは「1-Standard Code」	Annex I, Section 9の翻訳テーブルに対応する番号
transl_except	「transl_table」で定義されている遺伝暗号に適合しないコドンの翻訳を示す	(pos:<位置>, aa:<アミノ酸>)
codon_start	最初の塩基を基準としたCDSのリーディングフレームを示す	1, 2, または3

feature keyとqualifier

ヌクレオチド配列：「CDS」feature key

- 例 - SEQ ID NO: 1は、酵母ミトコンドリア遺伝子由来のコード配列（遺伝暗号表「3-Yeast Mitochondrial Code」）の断片です：

```
...tg gat aat gaa gaa gtt aac gaa gaa tgt atg aga tta ttt ttc aag aac gct cgt cat cta aca tca agg ttg aca taa      80
... Asp Asn Glu Glu Val Asn Glu Glu Cys Sec Arg Leu Phe Phe Lys Asn Ala Arg His Thr Thr Ser Arg Leu Thr *
```



30-32位のコドンはセレノシステイン (Sec) をコードする

配列はコドンの一部で始まる

- この配列を正確に表すために、CDS featureに情報を含めるべき情報とは？

feature key と qualifier

ヌクレオチド配列：「CDS」 feature key

- 例 - SEQ ID NO: 1は、酵母ミトコンドリア遺伝子由来のコード配列（遺伝暗号表「3-Yeast Mitochondrial Code」）の断片です：

```
...tg gat aat gaa gaa gtt aac gaa gaa tgt atg aga tta ttt ttc aag aac gct cgt cat cta aca tca agg ttg aca taa      80
... Asp Asn Glu Glu Val Asn Glu Glu Cys Sec Arg Leu Phe Phe Lys Asn Ala Arg His Thr Thr Ser Arg Leu Thr *
```

- 配列表に示される配列は、次の通りです：

```
tggataatga agaagttaac gaagaatgta tgagattatt ttcaagaac gctcgtcatc taacatcaag gttgacataa
```

feature keyとqualifier

ヌクレオチド配列：「CDS」 feature key

- 例 - SEQ ID NO: 1は、酵母ミトコンドリア遺伝子由来のコード配列（遺伝暗号表「3-Yeast Mitochondrial Code」）の断片です：

```
...tg gat aat gaa gaa gtt aac gaa gaa tgt atg aga tta ttt ttc aag aac gct cgt cat cta aca tca agg ttg aca taa      80
... Asp Asn Glu Glu Val Asn Glu Glu Cys Sec Arg Leu Phe Phe Lys Asn Ala Arg His Thr Thr Ser Arg Leu Thr *
```

- 「CDS」 feature key
- featureの位置：<1..80

「<」は、コード領域が1位より前から開始することを示す

この位置は78-80位に終止コドンを含む

feature key と qualifier

ヌクレオチド配列：「CDS」 feature key

- 例 - SEQ ID NO: 1は、酵母ミトコンドリア遺伝子由来のコード配列（遺伝暗号表「3-Yeast Mitochondrial Code」）の断片です：

```
...tg gat aat gaa gaa gtt aac gaa gaa tgt atg aga tta ttt ttc aag aac gct cgt cat cta aca tca agg ttg aca taa      80
... Asp Asn Glu Glu Val Asn Glu Glu Cys Sec Arg Leu Phe Phe Lys Asn Ala Arg His Thr Thr Ser Arg Leu Thr *
```

- 「CDS」 feature key
- featureの位置：<1..80
- 値が「3」の「codon_start」 qualifier

「codon_start」 qualifierの値の3は、最初の完全なコドンがこの位置内の3番目の位置から開始することを示す

feature key と qualifier

ヌクレオチド配列：「CDS」 feature key

- 例 - SEQ ID NO: 1は、酵母ミトコンドリア遺伝子由来のコード配列（遺伝暗号表「3-Yeast Mitochondrial Code」）の断片です：

```
...tg gat aat gaa gaa gtt aac gaa gaa tgt atg aga tta ttt ttc aag aac gct cgt cat cta aca tca agg ttg aca taa      80
... Asp Asn Glu Glu Val Asn Glu Glu Cys Sec Arg Leu Phe Phe Lys Asn Ala Arg His Thr Thr Ser Arg Leu Thr *
```

- 「CDS」 feature key
- featureの位置：<1..80
- 値が「3」の「codon_start」 qualifier
- 値が「3」の「transl_table」 qualifier

Annex 1, Section 9の
Genetic Code Tablesを使用
して、「transl_table」
qualifierの値を決定する。

feature key と qualifier

ヌクレオチド配列：「CDS」 feature key

- 例 - SEQ ID NO: 1は、酵母ミトコンドリア遺伝子由来のコード配列（遺伝暗号表「3-Yeast Mitochondrial Code」）の断片です：

```
...tg gat aat gaa gaa gtt aac gaa gaa tgt atg aga tta ttt ttc aag aac gct cgt cat cta aca tca agg ttg aca taa 80
... Asp Asn Glu Glu Val Asn Glu Glu Cys Sec Arg Leu Phe Phe Lys Asn Ala Arg His Thr Thr Ser Arg Leu Thr *
```

- 「CDS」 feature key
- featureの位置：<1..80
- 値が「3」の「codon_start」 qualifier
- 値が「3」の「transl_table」 qualifier
- 値が「(pos:30..32,aa:Sec)」の「transl_except」 qualifier

「transl_except」 qualifierは、
アミノ酸セレノシステインを
含む

feature key と qualifier

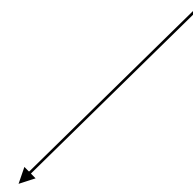
ヌクレオチド配列：「CDS」 feature key

- 例 - SEQ ID NO: 1は、酵母ミトコンドリア遺伝子由来のコード配列（遺伝暗号表「3-Yeast Mitochondrial Code」）の断片です：

```
...tg gat aat gaa gaa gtt aac gaa gaa tgt atg aga tta ttt ttc aag aac gct cgt cat cta aca tca agg ttg aca taa      80
... Asp Asn Glu Glu Val Asn Glu Glu Cys Sec Arg Leu Phe Phe Lys Asn Ala Arg His Thr Thr Ser Arg Leu Thr *
```

- 「CDS」 feature key
- featureの位置：<1..80
- 値が「3」の「codon_start」 qualifier
- 値が「3」の「transl_table」 qualifier
- 値が「(pos:30..32,aa:Sec)」の「transl_except」 qualifier
- 値が「**DNEEVNEECURLFFKNARHTTSRLT**」の「translation」 qualifier

終止コドンはtranslation
qualifierに含まれません！



feature key と qualifier

ヌクレオチド配列：「CDS」 feature key

- 例 - SEQ ID NO: 1は、酵母ミトコンドリア遺伝子由来のコード配列（遺伝暗号表「3-Yeast Mitochondrial Code」）の断片です：

```
...tg gat aat gaa gaa gtt aac gaa gaa tgt atg aga tta ttt ttc aag aac gct cgt cat cta aca tca agg ttg aca taa      80
... Asp Asn Glu Glu Val Asn Glu Glu Cys Sec Arg Leu Phe Phe Lys Asn Ala Arg His Thr Thr Ser Arg Leu Thr *
```

- 「CDS」 feature key
- featureの位置：<1..80
- 値が「3」の「codon_start」 qualifier
- 値が「3」の「transl_table」 qualifier
- 値が「(pos:30..32,aa:Sec)」の「transl_except」 qualifier
- 値が「DNEEVNEECURLFFKNARHTTSRLT」の「translation」 qualifier
- translationのための別個のタンパク質配列

feature key と qualifier

ヌクレオチド配列：「CDS」 feature key

- 例 - SEQ ID NO: 1は、酵母ミトコンドリア遺伝子由来のコード配列（遺伝暗号表「3-Yeast Mitochondrial Code」）の断片です：

```
...tg gat aat gaa gaa gtt aac gaa gaa tgt atg aga tta ttt ttc aag aac gct cgt cat cta aca tca agg ttg aca taa      80
... Asp Asn Glu Glu Val Asn Glu Glu Cys Sec Arg Leu Phe Phe Lys Asn Ala Arg His Thr Thr Ser Arg Leu Thr *
```

- 「CDS」 feature key
- featureの位置：<1..80
- 値が「3」の「codon_start」 qualifier
- 値が「3」の「transl_table」 qualifier
- 値が「(pos:30..32,aa:Sec)」の「transl_except」 qualifier
- 値が「**DNEEVNEECURLFFKNARHTTSRLT**」の「translation」 qualifier
- translationのための新しいタンパク質配列
- 翻訳されたタンパク質のSEQ ID番号を持つ「protein_id」 qualifier

feature key と qualifier

ヌクレオチド配列：「CDS」 feature key

- 「CDS」 feature の位置では、「join」位置演算子を使用することで、配列の不連続なセグメントを結合して単一のコード領域にできます

join(位置1,位置2)

- 「CDS」 feature の位置では、「complement」演算子を使用して、位置記述子で特定された配列に相補的な鎖にそのfeatureが位置することを示すことができます

complement(位置)

feature key と qualifier

アミノ酸配列：よく使用されるfeature key

Feature key	説明	必須のqualifier？
SITE	関心のある単一のアミノ酸位置を示す	必須のNOTE qualifierで、位置を記述しなければならない
REGION	関心のある領域を示す	なし；NOTEは任意
BINDING	化学基の結合部位を示す	必須のNOTE qualifierは、化学基の名前を含んでいなければならない
UNSURE	配列内の不確かな領域を記述する	なし；NOTEは任意

feature key と qualifier

アミノ酸配列：修飾アミノ酸

- 「修飾アミノ酸」とは、次のものを除くアミノ酸です：

L-アラニン	L-アルギニン	L-アスパラギン	L-アスパラギン酸
L-システイン	L-グルタミン	L-グルタミン酸	L-グリシン
L-ヒスチジン	L-イソロイシン	L-ロイシン	L-リシン
L-メチオニン	L-フェニルアラニン	L-プロリン	L-ピロリシン
L-セリン	L-セレノシステイン	L-スレオニン	L-トリプトファン
	L-チロシン	L-バリン	

(ST.26, pgh.3(e))

- 「修飾アミノ酸」は、可能な限り、対応する未修飾アミノ酸で表す必要があります。それ以外の場合は「X」で表すことができます。例えば、「ヒドロキシリジン」は配列中、「K」とします。「オルニチン」は「X」とします。(ST.26, pgh.29)
- 記号「X」は、1つの残基のみに対応します。

feature keyとqualifier

アミノ酸配列：修飾アミノ酸

- いくつかのfeature keyを使用して「修飾アミノ酸」を示すことができます：

Feature key	説明
SITE	翻訳後修飾されていないアミノ酸を示す
MOD_RES	翻訳後修飾されたアミノ酸を示す
CARBOHYD	グリコシル化アミノ酸を示す
LIPID	アミノ酸に脂質部分が共有結合していることを示す

- 上記のfeature keyのそれぞれは、修飾を記述する値を持つ必須のNOTE qualifierを含む必要があります。

feature key と qualifier

アミノ酸配列：修飾アミノ酸

Gly-Ser-N-acetylAla-Ser-Asp-Val-Orn-Lys-Asn-Val-Leu
 1 5 10

3位のアラニンは
細胞内で翻訳後修飾されてN-アセチルアラニンになる

シアリルオリゴ糖

- 配列表に示される配列は、次の通りです：

GSASDVXKNVL

- 位置が「3」の「MOD_RES」 feature key および値が「N-acetylalanine」のNOTE qualifier

feature key と qualifier

アミノ酸配列：修飾アミノ酸

Gly-Ser-N-acetylAla-Ser-Asp-Val-Orn-Lys-Asn-Val-Leu
 1 5 10

3位のアラニンは
細胞内で翻訳後修飾されてN-アセチルアラニンになる

シアリルオリゴ糖

- 配列表に示される配列は、次の通りです：

GSASDVXKNVL

- 位置が「3」の「MOD_RES」 feature key および値が「N-acetylalanine」のNOTE qualifier
- 位置が「7」の「SITE」 feature key および値が「ornithine」のNOTE qualifier

featureの位置の形式

位置の形式

すべての分子型の位置記述子

- 位置記述子は、配列中のfeatureの位置を識別するために使用されます
- ST.26には、位置記述子の形式に関する必須要件があります
- ヌクレオチドおよびアミノ酸配列には、次の位置記述子の形式を使用できます :

Location descriptor type	Syntax	Description
Single residue number	x	Points to a single residue in the sequence.
Residue numbers delimitating a sequence span	x..y	Points to a continuous range of residues bounded by and including the starting and ending residues.
Residues before the first or beyond the last specified residue number	<x >x <x..y x..>y <x..>y	Points to a region including a specified residue or span of residues and extending beyond a specified residue. The '<' and '>' symbols may be used with a single residue or the starting and ending residue numbers of a span of residues to indicate that a feature extends beyond the specified residue number.

WIPO Standard ST.26, paragraph 66(a)

位置の形式

すべての分子型の位置記述子

Location Example	Description
467	Points to residue 467 in the sequence.
340..565	Points to a continuous range of residues bounded by and including residues 340 and 565.
<1	Points to a feature location before the first residue.
<345..500	Indicates that the exact lower boundary point of a feature is unknown. The location begins at some residue previous to 345 and continues to and includes residue 500.
<1..888	Indicates that the feature starts before the first sequenced residue and continues to and includes residue 888.
1..>888	Indicates that the feature starts at the first sequenced residue and continues beyond residue 888.
<1..>888	Indicates that the feature starts before the first sequenced residue and continues beyond residue 888.

WIPO Standard ST.26, paragraph 70(a)

位置の形式

ヌクレオチド配列限定の位置記述子

- 次の位置記述子の形式は、DNAおよびRNA配列にのみ使用できます。

Location descriptor type	Syntax	Description
A site between two adjoining nucleotides	x^y	Points to a site between two adjoining nucleotides, e.g., endonucleolytic cleavage site. The position numbers for the adjacent nucleotides are separated by a caret (^). The permitted formats for this descriptor are x^{x+1} (for example 55 ⁵⁶), or, for circular nucleotides, x^1 , where “x” is the full length of the molecule, i.e. 1000 ¹ for circular molecule with length 1000.

WIPO Standard ST.26, paragraph 66(b)

位置の形式

ヌクレオチド配列限定の位置記述子

- x^y 位置形式を使用する場合、 x と y は隣接する残基でなければなりません

Location example	Description
123 ¹²⁴	Points to a site between residues 123 and 124.
867 ¹	In a circular molecule with 867 residues, points to a site between the residue indicated as position 1 and the residue indicated as position 867

WIPO Standard ST.26, paragraph 70(b)

位置の形式

位置記述子 – アミノ酸配列の特殊なケース

- 位置記述子 $x..y$ は、「CROSSLNK」または「DISULFID」 feature keyとともに使用された場合、示された2つの残基間の鎖内架橋を示します

Location descriptor type	Syntax	Description
Residue numbers joined by an intrachain cross-link	$x..y$	Points to amino acids joined by an intrachain linkage when used with a feature that indicates an intrachain cross-link, such as “CROSSLNK” or “DISULFID”.

WIPO Standard ST.26, paragraph 66(c)

位置の形式

位置記述子 – アミノ酸配列の特殊なケース

```

<SequenceData sequenceIDNumber="4">
  <INSDSeq>
    <INSDSeq_length>81</INSDSeq_length>
    <INSDSeq_moltype>AA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>SOURCE</INSDFeature_key>
        <INSDFeature_location>1..81</INSDFeature_location>
        <INSDFeature_qual>
          <INSDQualifier>
            <INSDQualifier_name>MOL_TYPE</INSDQualifier_name>
            <INSDQualifier_value>protein</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q25">
            <INSDQualifier_name>ORGANISM</INSDQualifier_name>
            <INSDQualifier_value>synthetic construct</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_qual>
      </INSDFeature>
      <INSDFeature>
        <INSDFeature_key>DISULFID</INSDFeature_key>
        <INSDFeature_location>30..50</INSDFeature_location>
        <INSDFeature_qual>
          <INSDQualifier id="q22">
            <INSDQualifier_name>NOTE</INSDQualifier_name>
            <INSDQualifier_value>disulfide bond</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_qual>
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>QKKMIQFFKITHRYYYDIIIEHLCAKYDMNSVISNALFAKLNLMQYTDGLSTHEKIIINTSNPLTFSIVISLQRCVINLGGST</INSDSeq_sequence>
  </INSDSeq>
</SequenceData>

```


位置の形式

ヌクレオチド配列の位置演算子

- DNAおよびRNA配列では、「join」、「order」、「complement」の3つの位置演算子が利用可能です
- join位置演算子の使用は、位置記述子によって記述されるヌクレオチド残基が生物学的プロセスによって物理的に接触することを意味します（ST.26, paragraph 68）
- 「join」と「order」を使用する位置には、少なくとも2つのコンマで区切られた位置記述子が必要です
- 「complement」は「join」または「order」と組み合わせて使用できます

Location syntax	Location description
<code>join(location, location, ..., location)</code>	The indicated locations are joined (placed end-to-end) to form one contiguous sequence.
<code>order(location, location, ..., location)</code>	The elements are found in the specified order but nothing is implied about whether joining those elements is reasonable.
<code>complement(location)</code>	Indicates that the feature is located on the strand complementary to the sequence span specified by the location descriptor, when read in the 5' to 3' direction or in the direction that mimics the 5' to 3' direction.

位置の形式

ヌクレオチド配列の位置演算子

Location example	Description
<code>join(12..78,134..202)</code>	Indicates that regions 12 to 78 and 134 to 202 should be joined to form one contiguous sequence.
<code>order(15..228,341..502)</code>	Indicates that regions 15 to 228 and 341 to 502 are present in the specified order
<code>complement(34..126)</code>	Starts at the nucleotide complementary to 126 and finishes at the nucleotide complementary to nucleotide 34 (the feature is on the strand complementary to the presented strand).
<code>complement(join(2691..4571,4918..5163))</code>	Joins nucleotides 2691 to 4571 and 4918 to 5163, then complements the joined segments (the feature is on the strand complementary to the presented strand).
<code>join(complement(4918..5163),complement(2691..4571))</code>	Complements regions 4918 to 5163 and 2691 to 4571, then joins the complemented segments (the feature is on the strand complementary to the presented strand).

WIPO Standard ST.26, paragraph 70(b)

質疑応答

Qualifierの値の形式

Qualifierの値

- Qualifierはfeatureをさらに定義します

Qualifierの値

- Qualifierはfeatureをさらに定義します
- Qualifierは、qualifierの名前と、多くの場合、qualifierの値とで構成されます

```
<INSDQualifier>  
  <INSDQualifier_name>transl_table</INSDQualifier_name>  
  <INSDQualifier_value>12</INSDQualifier_value>  
</INSDQualifier>
```

Qualifierの値

- Qualifierはfeatureをさらに定義します
- Qualifierは、qualifierの名前と、多くの場合、qualifierの値とで構成されます

```
<INSDQualifier>  
  <INSDQualifier_name>transl_table</INSDQualifier_name>  
  <INSDQualifier_value>12</INSDQualifier_value>  
</INSDQualifier>
```
- 各feature keyには、そのfeatureに使用可能なqualifierのリストがあります。一部のfeature keyには、必須のqualifierがあります。

Qualifierの値

- Qualifierはfeatureをさらに定義します
- Qualifierは、qualifierの名前と、多くの場合、qualifierの値とで構成されます

```
<INSDQualifier>  
  <INSDQualifier_name>transl_table</INSDQualifier_name>  
  <INSDQualifier_value>12</INSDQualifier_value>  
</INSDQualifier>
```
- 各feature keyには、そのfeatureに使用可能なqualifierのリストがあります。一部のfeature keyには、必須のqualifierがあります。
- feature keyとその使用可能なqualifierは、ST.26, Annex IのSection 5（ヌクレオチド配列）およびSection 7（アミノ酸配列）に記載されています。

Qualifierの値

- 例 - feature key 「misc_binding」 には、1つの必須のqualifier 「bound_moiety」と6つの任意のqualifierがあります：

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

(ST.26, Annex I, Section 5.12)

Qualifierの値

- Qualifierはfeatureをさらに定義します
- Qualifierは、qualifierの名前と、多くの場合、qualifierの値とで構成されます

```
<INSDQualifier>  
  <INSDQualifier_name>transl_table</INSDQualifier_name>  
  <INSDQualifier_value>12</INSDQualifier_value>  
</INSDQualifier>
```
- 各feature keyには、そのfeatureに使用可能なqualifierのリストがあります。一部のfeature keyには、必須のqualifierがあります。
- feature keyとその使用可能なqualifierは、ST.26, Annex IのSection 5（ヌクレオチド配列）およびSection 7（アミノ酸配列）に記載されています。
- ヌクレオチド配列のqualifierは、その説明とともにST.26, Annex IのSection 6（ヌクレオチド配列）およびSection 8（アミノ酸配列）に記載されています。

Qualifierの値

6.3. Qualifier	bound_moiety
Definition	name of the molecule/complex that may bind to the given feature
Mandatory value format	free text Language-dependent: this value may require translation for National/Regional procedures
Example	<INSDQualifier_value>GAL4</INSDQualifier_value>
Comment	A single bound_moiety qualifier is permitted on the "misc_binding", "oriT" and "protein_bind" features.

6.22. Qualifier	gene
Definition	symbol of the gene corresponding to a sequence region
Mandatory value format	free text
Example	<INSDQualifier_value>ilvE</INSDQualifier_value>
Comment	Use gene qualifier to provide the gene symbol; use standard_name qualifier to provide the full gene name.

(ST.26, Annex I, Section 6.3および6.22)

Qualifierの値

形式タイプ

■ qualifierの値には、いくつかの形式タイプがあります：

1. 事前に定義された値の選択肢を持つqualifier
2. 定義された値の形式を持つqualifier
3. 値が配列であるqualifier
4. 値のないqualifier
5. 「フリーテキスト」値を持つqualifier
 - 「フリーテキスト」 qualifierの値のサブセットは「言語依存」に分類されます

Qualifierの値

形式タイプ – 事前に定義された値の選択肢

- 事前に定義された値の選択肢を持つqualifier

- 例：

「codon_start」 – 値は「1」、「2」、または「3」が可能

6.9.	Qualifier	codon_start
	Definition	indicates the offset at which the first complete codon of a coding feature can be found, relative to the first base of that feature.
	Mandatory value format	1 or 2 or 3
	Example	<INSDQualifier_value>2</INSDQualifier_value>

Qualifierの値

書式タイプ – 事前に定義された値の選択肢

- 例：「rpt_type」には限定された値の選択肢があります

6.60. Qualifier	rpt_type
Definition	structure and distribution of repeated sequence
Mandatory value format	One of the following controlled vocabulary terms or phrases: tandem direct inverted flanking nested terminal dispersed long_terminal_repeat non_ltr_retrotransposon_polymeric_tract centromeric_repeat telomeric_repeat x_element_combinatorial_repeat y_prime_element other
Example	<INSDQualifier_value>inverted</INSDQualifier_value> <INSDQualifier_value>long_terminal_repeat</INSDQualifier_value>
Comment	Definitions of the values: tandem - a repeat that exists adjacent to another in the same orientation; direct - a repeat that exists not always adjacent but is in the same orientation; inverted - a repeat pair occurring in reverse orientation to one another on the same molecule;

Qualifierの値

書式タイプ – 事前に定義された値の選択肢

- 事前に定義された値の選択肢を持つqualifierの場合、WIPO Sequenceは、使用できるすべての値を事前に設定されたドロップダウンリストに表示します：

The screenshot shows the 'QUALIFIERS' section of the WIPO Sequence interface. A dropdown menu is open for the 'Qualifier Value' field, listing various options. The 'dispersed' option is currently selected and highlighted in blue. Below the dropdown, the 'SEQUENCE' section displays a grid of sequence data.

Qualifier Name *	Qualifier Value
rpt_type	<ul style="list-style-type: none"> tandem direct inverted flanking nested terminal dispersed long_terminal_repeat non_ltr_retrotransposon_polymeric_tract centromeric_repeat telomeric_repeat x_element_combinatorial_repeat y_prime_element other

SEQUENCE

tgcgatatt	atcataagc	atgggcgtct	ggaagtgcgc	ctgacggccc	agaagggcgt
cttactgtta	caagaagagt	atcatgacgc	gcatcatagc	cacttagacc	cgcacttatg
gttggcgcgc	caaaatatca	ttgccctcgt	gcaaggattg	gataaacata	ctgctgagtt

Qualifierの値

形式タイプ – 定義された値の形式

- 定義された値の形式を持つ qualifier
- 例：「anticodon」 – 値は次の形式でなければなりません
「(pos:<位置>,aa:<アミノ酸>,seq:<テキスト>)」

6.2.	Qualifier	anticodon
	Definition	location of the anticodon of tRNA and the amino acid for which it codes
	Mandatory value format	(pos:<location>,aa:<amino_acid>,seq:<text>) where <location> is the position of the anticodon and <amino_acid> is the three letter abbreviation for the amino acid encoded and <text> is the sequence of the anticodon
	Example	<pre><INSDQualifier_value>(pos:34..36,aa:Phe,seq:aaa)</INSDQualifier_value> <INSDQualifier_value>(pos:join(5,495..496),aa:Leu,seq:taa)</INSDQualifier_value> <INSDQualifier_value>(pos:complement(4156..4158),aa:Glu,seq:ttg)</INSDQualifier_value></pre>

Qualifierの値

形式タイプ - 配列

- 値が配列であるqualifier
- 例：「translation」 – 値は1文字のアミノ酸略号を使用した配列でなければなりません

6.79. Qualifier	translation
Definition	one-letter abbreviated amino acid sequence derived from either the standard (or universal) genetic code or the table as specified in a transl_table qualifier and as determined by an exception in the transl_except qualifier
Mandatory value format	contiguous string of one-letter amino acid abbreviations from Section 3 of this Annex, "X" is to be used for AA exceptions.
Example	<INSDQualifier_value>MASTFPPWYRGCASTPSLKGLIMCTW</INSDQualifier_value>
Comment	to be used with CDS feature only; must be accompanied by protein_id qualifier when the translation product contains four or more specifically defined amino acids; see transl_table for definition and location of genetic code Tables; only one of the qualifiers translation, pseudo and pseudogene are permitted to further annotate a CDS feature.

Qualifierの値

形式のタイプ - 配列

- qualifier 「replace」 – 値は、単一のヌクレオチド残基、残基の配列、または空にすることができます
- 「replace」の空の値は、対応するfeatureに示されている残基の欠失を意味します

6.57. Qualifier	replace
Definition	indicates that the sequence identified in a feature's location is replaced by the sequence shown in the qualifier's value; if no sequence (i.e., no value) is contained within the qualifier, this indicates a deletion
Mandatory value format	free text
Example	<pre><INSDQualifier_value>a</INSDQualifier_value> <INSDQualifier_value></INSDQualifier_value> - for a deletion</pre>

Qualifierの値

形式のタイプ - 配列

- qualifier 「replace」 – 値は、単一のヌクレオチド残基、残基の配列、または空にすることができます
- 「replace」の空の値は、対応するfeatureに示されている残基の欠失を意味します

6.57. Qualifier	replace
Definition	indicates that the sequence identified in a feature's location is replaced by the sequence shown in the qualifier's value; if no sequence (i.e., no value) is contained within the qualifier, this indicates a deletion
Mandatory value format	free text
Example	<pre><INSDQualifier value>a</INSDQualifier value> <INSDQualifier_value></INSDQualifier_value> - for a deletion</pre>

「空の値」

Qualifierの値

形式タイプ - 値なし

- 値のないqualifier

- 例：

“environmental_sample” “germline” “macronuclear” “proviral”

6.51. Qualifier	proviral
Definition	this qualifier is used to flag sequence obtained from a virus or phage that is integrated into the genome of another organism
Value format	none

- WIPO Sequenceでは、値のないqualifierに値を追加できません

- これらのqualifierは、空の「INSDQualifier_value」要素を持つことはできません

Qualifierの値

形式タイプ - フリーテキスト

- 多くのqualifierは「フリーテキスト」の値の形式を持ちます
- ST.26のparagraph 3(n)に、「フリーテキスト」は「... 特定のqualifierのための値の形式の一種であり、記述的なテキストフレーズの形式または他の指定された形式 (Annex Iに示す) で表されます」と定義されています。
- フリーテキストqualifierの値は1 0 0 0文字に制限されています (ST.26, paragraph 86)
- 「フリーテキスト」の値の形式を持つqualifierのサブセットは、「言語依存」です
- 「言語依存」フリーテキストは「国内、特定地域、または国際的な手続きでは、翻訳が必要な場合があります。」 ST.26, paragraph 3(o)

Qualifierの値

形式のタイプ - フリーテキスト

- 「言語依存フリーテキスト」値を持つqualifierは、翻訳が必要な場合があります
- フリーテキストの値の形式を持つqualifierが言語に依存するかどうかをどのように判断するか？
 - ST.26, Annex I, Section 6のTable 5に、言語依存フリーテキスト値の形式を持つすべてのヌクレオチド配列qualifierが記載されています
 - ST.26, Annex I, Section 8のTable 6に、言語依存フリーテキスト値の形式を持つすべてのアミノ酸配列qualifierが記載されています
 - qualifierの説明の「必須の値の形式」を見てください

6.5. Qualifier	cell_type
Definition	cell type from which the sequence was obtained
Mandatory value format	free text Language-dependent: this value may require translation for National/Regional procedures
Example	<INSDQualifier_value>leukocyte</INSDQualifier_value>

Qualifierの値

形式のタイプ - フリーテキスト

6.20. Qualifier	frequency
Definition	frequency of the occurrence of a feature
Mandatory value format	free text representing the proportion of a population carrying the feature expressed as a fraction
Example	<pre><INSDQualifier_value>23/108</INSDQualifier_value> <INSDQualifier_value>1 in 12</INSDQualifier_value> <INSDQualifier_value>0.85</INSDQualifier_value></pre>
6.21. Qualifier	function
Definition	function attributed to a sequence
Mandatory value format	free text Language-dependent: this value may require translation for National/Regional procedures
Example	<pre><INSDQualifier_value>essential for recognition of cofactor </INSDQualifier_value></pre>
Comment	The function qualifier is used when the gene name and/or product name do not convey the function attributable to a sequence.

Qualifierの値

形式タイプ – 「言語依存フリーテキスト」

- 「言語依存フリーテキスト」 qualifierの値は、配列表XMLに2つの言語（英語と1つの英語以外の言語）で記述できます（ST.26, paragraph 87）

Qualifierの値

形式の種類 – 「言語依存フリーテキスト」

- 「言語依存フリーテキスト」 qualifierの値は、配列表XMLに2つの言語（英語と英語以外の言語）で記述できます（ST.26, paragraph 87）
- 英語の言語依存qualifierの値は、INSDQualifier_value要素に記述する必要があります

Qualifierの値

形式の種類 – 「言語依存フリーテキスト」

- 「言語依存フリーテキスト」 qualifierの値は、配列表XMLに2つの言語（英語と英語以外の言語）で記述できます（ST.26, paragraph 87）
- 英語の言語依存qualifierの値は、INSDQualifier_value要素に記述する必要があります
- 英語以外の言語での言語依存qualifierの値は、NonEnglishQualifier_value要素に記述する必要があります

Qualifierの値

形式の種類 – 「言語依存フリーテキスト」

- 「言語依存フリーテキスト」 qualifierの値は、配列表XMLに2つの言語（英語と英語以外の言語）で記述できます（ST.26, paragraph 87）
- 英語の言語依存qualifierの値は、INSDQualifier_value要素に記述する必要があります
- 英語以外の言語での言語依存qualifierの値は、NonEnglishQualifier_value要素に記述する必要があります
- NonEnglishQualifier_value要素は、言語依存フリーテキスト値の形式を持つqualifierに対してのみ使用可能です（ST.26, paragraph 87 (b)）

Qualifierの値

形式の種類 – 「言語依存フリーテキスト」

```

<SequenceData sequenceIDNumber="2">
  <INSDSeq>
    <INSDSeq_length>29</INSDSeq_length>
    <INSDSeq_moltype>AA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>SOURCE</INSDFeature_key>
        <INSDFeature_location>1..29</INSDFeature_location>
        <INSDFeature_qual>
          <INSDQualifier>
            <INSDQualifier_name>ORGANISM</INSDQualifier_name>
            <INSDQualifier_value>synthetic construct</INSDQualifier_value>
          </INSDQualifier>
          - <INSDQualifier>
            <INSDQualifier_name>MOL_TYPE</INSDQualifier_name>
            <INSDQualifier_value>protein</INSDQualifier_value>
          </INSDQualifier>
          - <INSDQualifier id="q1">
            <INSDQualifier_name>NOTE</INSDQualifier_name>
            <INSDQualifier_value>Synthetic peptide antigen fragment</INSDQualifier_value>
            <NonEnglishQualifier_value>Synthetisches Peptidantigenfragment</NonEnglishQualifier_value>
          </INSDQualifier>
        </INSDFeature_qual>
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>GSLSDVRKDVVEKRIDKALEAFKNKMDKEK</INSDSeq_sequence>
  </INSDSeq>
</SequenceData>
<SequenceData sequenceIDNumber="3">

```

Qualifierの値

形式の種類 – 「言語依存フリーテキスト」

■ XMLの配列表に英語以外のqualifierの値が含まれている場合：

1. 配列表XMLのルート要素は、適切な2文字の言語コードの略語を持つ「nonEnglishFreeTextLanguageCode」属性を含んでいなければなりません（ST.26 paragraph 43 および 87 (b)）；

```
<ST26SequenceListing dtdVersion="V1_3" fileName="st26-annex-iii-sequence-listing-specimen.xml" softwareName="WIPO  
Sequence" softwareVersion="1.0" productionDate="2022-01-01" originalFreeTextLanguageCode="de"  
nonEnglishFreeTextLanguageCode="de">
```

Qualifierの値

形式の種類 – 「言語依存フリーテキスト」

■ XMLの配列表に英語以外のqualifierの値が含まれている場合：

1. 配列表XMLのルート要素は、適切な2文字の言語コードの略語を持つ「nonEnglishFreeTextLanguageCode」属性を含んでいなければなりません（ST.26 paragraph 43 および 87 (b)）；
2. 配列表内のすべての言語依存qualifierは、「nonEnglishFreeTextLanguageCode」属性で指定された言語の値を持たなければなりません；

Qualifierの値

形式の種類 – 「言語依存フリーテキスト」

■ XMLの配列表に英語以外のqualifierの値が含まれている場合：

1. 配列表XMLのルート要素は、適切な2文字の言語コードの略語を持つ「nonEnglishFreeTextLanguageCode」属性を含んでいなければなりません（ST.26 paragraph 43 および 87 (b)）；
2. 配列表内のすべての言語依存qualifierは、「nonEnglishFreeTextLanguageCode」属性で指定された言語の値を持たなければなりません；
3. NonEnglishQualifier_valueおよびINSDQualifier_valueの両方が単一のqualifierに対して存在する場合、2つの要素に含まれる情報は対応するものでなければなりません（ST.26 paragraph 87 (c)）。

Qualifierの値

形式の種類 – 「言語依存フリーテキスト」

- INSDQualifier 「id」属性とは？
- ST.26 paragraph 87(d)には、「言語依存qualifierの場合、INSDQualifier要素に任意の属性IDを含めることができます。この属性の値は、「q」の後に正の整数が続く形式でなければならず（例えば「q23」）、1つのINSDQualifier要素に対して一意でなければなりません。つまり、属性の値は、配列表ファイルで1度だけ使用されるものでなければなりません。」と記載されています。

```
<INSDQualifier id="q2">  
  <INSDQualifier_name>note</INSDQualifier_name>  
  <INSDQualifier_value>common name: tomato</INSDQualifier_value>  
  <NonEnglishQualifier_value>gemeinsamen Namen: Tomate</NonEnglishQualifier_value>  
</INSDQualifier>
```


Qualifierの値

形式の種類 – 「言語依存フリーテキスト」

- INSDQualifier 「id」 属性：
 - WIPO SequenceによるXLIFFファイルへのエクスポートに翻訳を必要とする可能性のあるqualifierの値を一意に識別する；
 - 任意；
 - 言語依存qualifierにのみ使用可能；
 - 配列表内で一意である必要がある；
 - WIPO Sequenceによって作成されるXMLに自動的に追加される。

Qualifierの値

形式の種類 – 「言語依存フリーテキスト」

■ 「originalFreeTextLanguageCode」属性

```
<ST26SequenceListing dtdVersion="V1_3" fileName="st26-annex-iii-sequence-listing-specimen.xml" softwareName="WIPO  
Sequence" softwareVersion="1.0" productionDate="2022-01-01" originalFreeTextLanguageCode="de"  
nonEnglishFreeTextLanguageCode="de">
```

- ST.26 paragraph 43に、「originalFreeTextLanguageCode」属性は、「言語依存フリーテキストqualifierが作成された単一の元の言語の言語コード...」と定義されています。
- この属性は任意です

特殊なケース：
DNA/RNAハイブリッド分子、
DNA中のウラシルおよび
RNA中のチミン

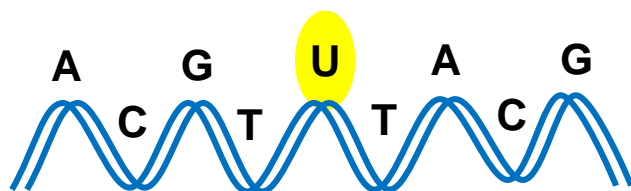
特殊なケース

- 注意：ST.26ではウラシルに「u」記号は使用できません
- DNAでは「t」はチミンです
- RNAでは「t」はウラシルです
- 考慮すべきシナリオは次の2つです：
 1. ウラシル核酸塩基を有するDNA分子またはチミン核酸塩基を有するRNA分子；
 2. DNA/RNAハイブリッド分子

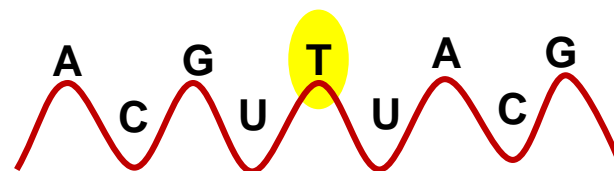
特殊なケース

DNA中のウラシルとRNA中のチミン

- 配列がウラシル核酸塩基を有するDNA骨格またはチミン核酸塩基を有するRNA骨格を有する場合 – 「修飾ヌクレオチド」として記述します



DNA



RNA

ST.26 Paragraph 14が適用されます :

「14. 記号「t」は、DNAではチミン、RNAではウラシルと解釈されます。DNA中のウラシルまたはRNA中のチミンは修飾ヌクレオチドとみなされ、paragraph 19に記載されるようにfeature tableに詳しく記述しなければなりません。」

特殊なケース

DNAおよびRNA配列

ある出願で次のRNA配列が開示されています：

5'-cgucccacgugtccgagqua-3'

- 12位の「チミン」に注意してください。この残基は、修飾ヌクレオチドとしての注釈付けが必要です。
- ST.26のparagraph 19には次のように記載されています：*DNA中のはウラシルまたはRNA中のチミンは、修飾ヌクレオチドとみなされます。配列では「t」として表され、feature key「modified_base」、qualifierの値として「OTHER」を持つqualifier「mod_base」、qualifierの値として「uracil」または「thymine」をそれぞれ持つqualifier「note」を用いてfeature tableにさらに記述する必要があります。*

特殊なケース

DNAおよびRNA配列

ある出願で次のRNA配列が開示されています：

5'-cgucccacgugtccgagqua-3'

- 12位の「チミン」に注意してください。この残基は、修飾ヌクレオチドとしての注釈付けが必要です。
- ST.26のparagraph 19には次のように記載されています：*DNA中のはウラシルまたはRNA中のチミンは、修飾ヌクレオチドとみなされます。配列では「t」として表され、feature key 「**modified_base**」、qualifier の値として「OTHER」を持つqualifier 「**mod_base**」、 「uracil」または「thymine」をそれぞれ持つqualifier 「**note**」を用いてfeature tableにさらに記述する必要があります。*

特殊なケース

DNAおよびRNA配列

ある出願で次のRNA配列が開示されています：

5'-cgucccacgugtccgaggua-3'

- ✓ すべてのウラシル残基は記号「t」で表さなければなりません。したがって、配列表で配列は次のように示されます：

cgtcccacgtgtccgaggta

特殊なケース

DNAおよびRNA配列

ある出願で次のRNA配列が開示されています：

5' -cgucccacgugtccgaggua -3'

- ✓ すべてのウラシル残基は記号「t」で表さなければなりません。したがって、配列表で配列は次のように示されます：

cgtccccacgtgtccgaggta

- ✓ 位置が「12」のfeature key「modified_base」

特殊なケース

DNAおよびRNA配列

ある出願で次のRNA配列が開示されています：

5' -cgucccacgugtccgaggua -3'

- ✓ すべてのウラシル残基は記号「t」で表さなければなりません。したがって、配列表で配列は次のように示されます：

cgtccccacgtgtccgaggta

- ✓ 位置が「12」のfeature key「modified_base」
- ✓ 値が「OTHER」のqualifier「mod_base」

特殊なケース

DNAおよびRNA配列

ある出願で次のRNA配列が開示されています：

5' -cgucccacgugtccgaggua -3'

- ✓ すべてのウラシル残基は記号「t」で表さなければなりません。したがって、配列表で配列は次のように示されます：

cgtccccacgtgtccgaggta

- ✓ 位置が「12」のfeature key「modified_base」
- ✓ 値が「OTHER」のqualifier「mod_base」
- ✓ 値が「thymine」のqualifier「note」

特殊なケース

DNAおよびRNA配列

```

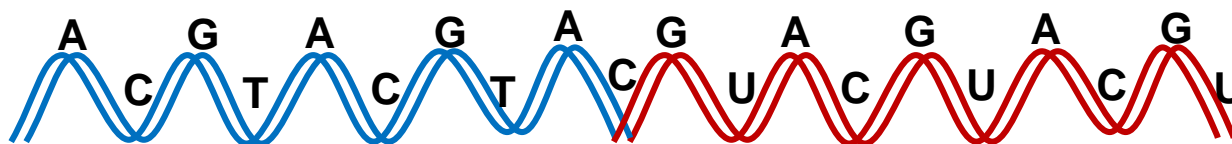
<SequenceData sequenceIDNumber="3">
  <INSDSeq>
    <INSDSeq_length>20</INSDSeq_length>
    <INSDSeq_moltype>RNA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..20</INSDFeature_location>
        <INSDFeature_qual>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>other RNA</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q7">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>synthetic construct</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_qual>
      </INSDFeature>
      <INSDFeature>
        <INSDFeature_key>modified_base</INSDFeature_key>
        <INSDFeature_location>12</INSDFeature_location>
        <INSDFeature_qual>
          <INSDQualifier>
            <INSDQualifier_name>mod_base</INSDQualifier_name>
            <INSDQualifier_value>OTHER</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q8">
            <INSDQualifier_name>note</INSDQualifier_name>
            <INSDQualifier_value>thymine</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_qual>
      </INSDFeature>
    </INSDSeq_feature-table>
    <INSDSeq_sequence>cgtcccacgtgtcgcgagta</INSDSeq_sequence>
  </INSDSeq>
</SequenceData>

```

特殊なケース

DNA/RNAハイブリッド配列

- 配列がハイブリッド分子である場合、すなわち、骨格の一部がDNAであり、骨格の一部がRNAである場合：



DNAセグメント RNAセグメント

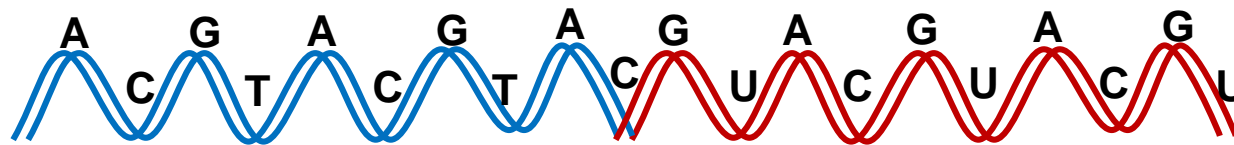
ST.26のParagraph 55が適用されます：

「55. 1つ以上のヌクレオチドのDNAセグメントとRNAセグメントの両方を含むヌクレオチド配列の場合、分子型はDNAとしなければなりません。結合したDNA/RNA分子については、feature key 「source」と、値が「synthetic construct」の必須のqualifier 「organism」および値が「other DNA」の必須のqualifier 「mol_type」とを用いてfeature tableにさらに記述する必要があります。結合したDNA/RNA分子のDNAおよびRNAセグメントのそれぞれについては、feature key 「misc_feature」と、セグメントがDNAかRNAかを示すqualifier 「note」を用いてさらに記述する必要があります。」

特殊なケース

DNA/RNAハイブリッド配列

- 配列がハイブリッド分子である場合、すなわち、骨格の一部がDNAであり、骨格の一部がRNAである場合：



DNAセグメント RNAセグメント

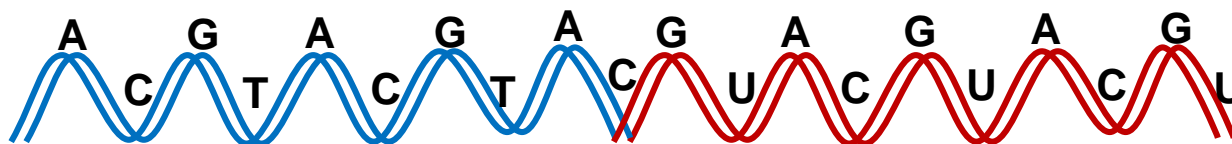
ST.26のParagraph 55が適用されます：

「55. 1つ以上のヌクレオチドのDNAセグメントとRNAセグメントの両方を含むヌクレオチド配列の場合、**分子型はDNAとしなければなりません**。結合したDNA/RNA分子については、feature key 「source」と、値が「synthetic construct」の必須のqualifier 「organism」および値が「other DNA」の必須のqualifier 「mol_type」とを用いてfeature tableにさらに記述する必要があります。結合したDNA/RNA分子のDNAおよびRNAセグメントのそれぞれについては、feature key 「misc_feature」と、セグメントがDNAかRNAかを示すqualifier 「note」を用いてさらに記述する必要があります。」

特殊なケース

DNA/RNAハイブリッド配列

- 配列がハイブリッド分子である場合、すなわち、骨格の一部がDNAであり、骨格の一部がRNAである場合：



DNAセグメント RNAセグメント

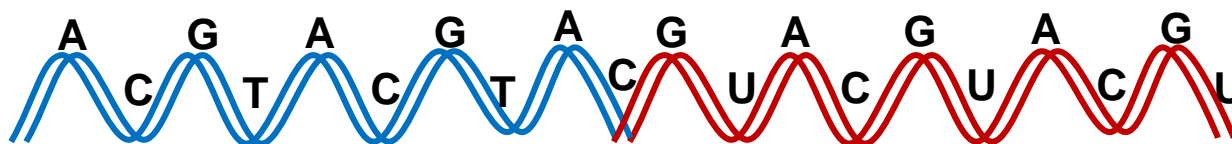
ST.26のParagraph 55が適用されます：

「55. 1つ以上のヌクレオチドのDNAセグメントとRNAセグメントの両方を含むヌクレオチド配列の場合、分子型はDNAとしなければなりません。結合したDNA/RNA分子については、feature key 「source」と、**値が「synthetic construct」の必須のqualifier 「organism」** および値が「other DNA」の必須のqualifier 「mol_type」とを用いてfeature tableにさらに記述する必要があります。結合したDNA/RNA分子のDNAおよびRNAセグメントのそれぞれについては、feature key 「misc_feature」と、セグメントがDNAかRNAかを示すqualifier 「note」を用いてさらに記述する必要があります。」

特殊なケース

DNA/RNAハイブリッド配列

- 配列がハイブリッド分子である場合、すなわち、骨格の一部がDNAであり、骨格の一部がRNAである場合：



DNAセグメント RNAセグメント

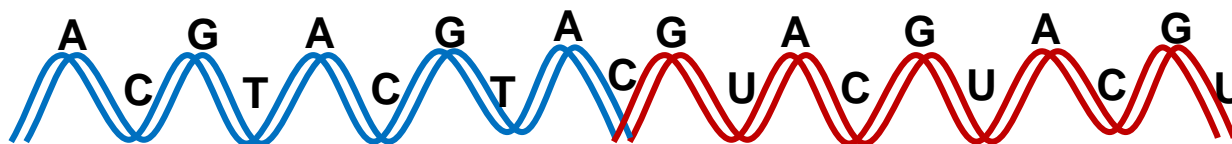
ST.26のParagraph 55が適用されます：

「55. 1つ以上のヌクレオチドのDNAセグメントとRNAセグメントの両方を含むヌクレオチド配列の場合、分子型はDNAとしなければなりません。結合したDNA/RNA分子については、feature key 「source」と、値が「synthetic construct」の必須のqualifier 「organism」および値が「other DNA」の必須のqualifier 「mol_type」とを用いてfeature tableにさらに記述する必要があります。結合したDNA/RNA分子のDNAおよびRNAセグメントのそれぞれについては、feature key 「misc_feature」と、セグメントがDNAかRNAかを示すqualifier 「note」を用いてさらに記述する必要があります。」

特殊なケース

DNA/RNAハイブリッド配列

- 配列がハイブリッド分子である場合、すなわち、骨格の一部がDNAであり、骨格の一部がRNAである場合：



DNAセグメント RNAセグメント

ST.26のParagraph 55が適用されます：

「55. 1つ以上のヌクレオチドのDNAセグメントとRNAセグメントの両方を含むヌクレオチド配列の場合、分子型はDNAとしなければなりません。結合したDNA/RNA分子については、feature key 「source」と、値が「synthetic construct」の必須のqualifier 「organism」および値が「other DNA」の必須のqualifier 「mol_type」とを用いてfeature tableにさらに記述する必要があります。結合したDNA/RNA分子のDNAおよびRNAセグメントのそれぞれについては、feature key 「misc_feature」と、セグメントがDNAかRNAかを示すqualifier 「note」を用いてさらに記述する必要があります。」

特殊なケース

DNA/RNAハイブリッド配列

ある出願で次のDNA/RNAハイブリッド配列が開示されています：

5' -ACCTGCcgucccacguguccgagguaGCATTA-3'

大文字の記号はDNA部分を表し、小文字の記号はRNA部分を表す。

- 残基1-6と27-32はDNA、残基7-26はRNA

- 次のことを考慮する：
 1. 生物の名称
 2. 分子型とmol_type
 3. DNAおよびRNAセグメントの識別

特殊なケース

DNAおよびRNA配列

ある出願で次のDNA/RNAハイブリッド配列が開示されています：

5' -ACCTGCcgucccacguguccgagguaGCATTA-3'

大文字の記号はDNA部分を表し、小文字の記号はRNA部分を表す。

- ST.26のParagraph 55には次のように記載されています：「...分子型はDNAとしなければなりません。結合したDNA/RNA分子については、feature key 「source」と、値が「synthetic construct」の必須のqualifier 「organism」および値が「other DNA」の必須のqualifier 「mol_type」とを用いてfeature tableにさらに記述する必要があります。」

- ✓ 分子型 = 「DNA」
- ✓ qualifier 「organism name」 = 「synthetic construct」
- ✓ qualifier 「mol_type」 = 「other DNA」

特殊なケース

DNAおよびRNA配列

ある出願で次のDNA/RNAハイブリッド配列が開示されています：

5' -ACCTGCcgucccacguguccgagguaGCATTA-3'

大文字の記号はDNA部分を表し、小文字の記号はRNA部分を表す。

- ✓ すべてのウラシル残基は記号「t」で表さなければなりません。したがって、配列表で配列は次のように示されます：

acctgccgtcccacgtgtccgaggtagcatta

特殊なケース

DNAおよびRNA配列

ある出願で次のDNA/RNAハイブリッド配列が開示されています：

5' -ACCTGCcgucccacguguccgagguaGCATTA-3'

大文字の記号はDNA部分を表し、小文字の記号はRNA部分を表す。

- ST.26のParagraph 55には次のように記載されています：「結合したDNA/RNA分子のDNAおよびRNAセグメントのそれぞれについては、feature key 「misc_feature」と、セグメントがDNAかRNAかを示すqualifier 「note」を用いてさらに記述する必要があります。」
- ✓ 3つのセグメント = 3つの「misc_feature」feature key

特殊なケース

DNAおよびRNA配列

ある出願で次のDNA/RNAハイブリッド配列が開示されています：

5' -ACCTGCcgucccacguguccgagguaGCATTA-3'

大文字の記号はDNA部分を表し、小文字の記号はRNA部分を表す。

- ✓ セグメント1、残基1-6：
位置が「1.. 6」の「misc_feature」feature key
値が「DNA」のqualifier「note」
- ✓ セグメント2、残基7-26：
位置が「7.. 26」の「misc_feature」feature key
値が「RNA」のqualifier「note」
- ✓ セグメント3、残基27-32
位置が「27.. 32」の「misc_feature」feature key
値が「DNA」のqualifier「note」

特殊なケース

DNAおよびRNA配列

```
<SequenceData sequenceIDNumber="4">
  <INSDSeq>
    <INSDSeq_length>32</INSDSeq_length>
    <INSDSeq_moltype>DNA</INSDSeq_moltype>
    <INSDSeq_division>PAT</INSDSeq_division>
    <INSDSeq_feature-table>
      <INSDFeature>
        <INSDFeature_key>source</INSDFeature_key>
        <INSDFeature_location>1..32</INSDFeature_location>
        <INSDFeature_quals>
          <INSDQualifier>
            <INSDQualifier_name>mol_type</INSDQualifier_name>
            <INSDQualifier_value>other DNA</INSDQualifier_value>
          </INSDQualifier>
          <INSDQualifier id="q10">
            <INSDQualifier_name>organism</INSDQualifier_name>
            <INSDQualifier_value>synthetic construct</INSDQualifier_value>
          </INSDQualifier>
        </INSDFeature_quals>
      </INSDFeature>
    </INSDSeq_feature-table>
  </INSDSeq>
</SequenceData>
```

特殊なケース

DNAおよびRNA配列

```

<INSDFeature>
  <INSDFeature_key>misc_feature</INSDFeature_key>
  <INSDFeature_location>1..6</INSDFeature_location>
  <INSDFeature_qual>
    <INSDQualifier id="q11">
      <INSDQualifier_name>note</INSDQualifier_name>
      <INSDQualifier_value>DNA</INSDQualifier_value>
    </INSDQualifier>
  </INSDFeature_qual>
</INSDFeature>
<INSDFeature>
  <INSDFeature_key>misc_feature</INSDFeature_key>
  <INSDFeature_location>7..26</INSDFeature_location>
  <INSDFeature_qual>
    <INSDQualifier id="q12">
      <INSDQualifier_name>note</INSDQualifier_name>
      <INSDQualifier_value>RNA</INSDQualifier_value>
    </INSDQualifier>
  </INSDFeature_qual>
</INSDFeature>
<INSDFeature>
  <INSDFeature_key>misc_feature</INSDFeature_key>
  <INSDFeature_location>27..32</INSDFeature_location>
  <INSDFeature_qual>
    <INSDQualifier id="q13">
      <INSDQualifier_name>note</INSDQualifier_name>
      <INSDQualifier_value>DNA</INSDQualifier_value>
    </INSDQualifier>
  </INSDFeature_qual>
</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>acctgccgtcccacgtgtccgaggtagcatta</INSDSeq_sequence>
</INSDSeq>
</SequenceData>
</ST26SequenceListing>

```

DNAセグメント 1

RNAセグメント 2

DNAセグメント 3

質疑応答

ヌクレオチド類似体、 D-アミノ酸、および分 岐配列

新しい分子型

ヌクレオチド類似体

- 1つ以上のヌクレオチド類似体を含む核酸配列は、ST.26の規定の対象です
- ヌクレオチド類似体は、ST.26のparagraph 3(g)(2)の「ヌクレオチド」の定義に包含されます：

「核酸類似体の骨格を形成する場合、2'デオキシリボース5'ーリン酸骨格またはリボース5'ーリン酸骨格を含む核酸中の核酸塩基の配列に類似した核酸塩基の配列が得られ、この核酸類似体は相補核酸と塩基対合することができる2'デオキシリボース5'ーリン酸またはリボース5'ーリン酸の類似体」

- 一般的なヌクレオチド類似体には、ペプチド核酸（PNA）、グリコール核酸（GNA）、トレオース核酸、モルフォリノなどがあります
- 5' -3'方向を示す、左から右方向に表さなければなりません。（ST.26, paragraph 11）

新しい分子型 ヌクレオチド類似体

ある特許出願で次のグリコール核酸（GNA）配列が開示されています：

PO_4 -tagttcattgactaaggctccccattgact-OH

配列の PO_4 末端はDNA配列の5'末端に対応する。

✓ この配列は、配列表に含めなければなりません

新しい分子型 ヌクレオチド類似体

ある特許出願で次のグリコール核酸（GNA）配列が開示されています：

PO_4 -tagttcattgactaaggctccccattgact-OH

配列の PO_4 末端はDNA配列の5'末端に対応する。

- ✓ この配列は、配列表に含めなければなりません
- ✓ PO_4 の末端は5'末端に対応するので、配列は示した方向に表さなければなりません

新しい分子型 ヌクレオチド類似体

ある特許出願で次のグリコール核酸（GNA）配列が開示されています：

PO_4 -tagttcattgactaaggctccccattgact-OH

配列の PO_4 末端はDNA配列の5'末端に対応する。

- ✓ この配列は、配列表に含めなければなりません
- ✓ PO_4 の末端は5'端に対応するので、配列は示した方向に表さなければなりません
- ✓ 配列全体に、「modified_base」feature key、値が「OTHER」の「mod_base」qualifier、および「glycol nucleic acids」などの修飾ヌクレオチドの省略されていない完全な名前を含むnote qualifierで注釈を付ける必要があります

新しい分子型

ヌクレオチド類似体

```

- <INSDFeature>
  <INSDFeature_key>modified_base</INSDFeature_key>
  <INSDFeature_location>1..30</INSDFeature_location>
  - <INSDFeature_qual>
    - <INSDQualifier>
      <INSDQualifier_name>mod_base</INSDQualifier_name>
      <INSDQualifier_value>OTHER</INSDQualifier_value>
    </INSDQualifier>
    - <INSDQualifier>
      <INSDQualifier_name>note</INSDQualifier_name>
      <INSDQualifier_value>2,3-dihydroxypropyl nucleosides (glycol nucleic acids)</INSDQualifier_value>
    </INSDQualifier>
  </INSDFeature_qual>
</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>tagttcattgactaaggctccccattgact</INSDSeq_sequence>
</INSDSeq>

```

注：この例のさらなる考察は
WIPO Standard ST.26, Annex VIのExample 3(g)-4にあります。

新しい分子型

D-アミノ酸

- 1つ以上のD-アミノ酸を含むアミノ酸配列は、ST.26の規定の対象です
- D-アミノ酸は、ST.26のparagraph 3 (a)の「アミノ酸」の定義に含まれています：

「「アミノ酸」とは、Annex Iに定める記号のいずれかを用いて表すことのできる任意のアミノ酸を意味します（Section 3, Table 3参照）。そのようなアミノ酸には、とりわけ、D-アミノ酸および修飾側鎖または合成側鎖を含むアミノ酸が含まれます。」

- D-アミノ酸は、配列中、可能な限り対応する未修飾L-アミノ酸記号として配列で表す必要があります
- feature tableに修飾アミノ酸として記述されなければなりません

新しい分子型

D-アミノ酸

ある特許出願で次の配列が開示されています：

D-Ala-D-Glu-Lys-Leu-Gly-D-Met

✓ この配列は、配列表に含めなければなりません

新しい分子型

D-アミノ酸

ある特許出願で次の配列が開示されています：

D-Ala-D-Glu-Lys-Leu-Gly-D-Met

- ✓ この配列は、配列表に含めなければなりません
- ✓ 次のように表します：AGKLGGM
- ✓ 1位のアラニン、2位のグルタミン酸、6位のメチオニンには、それぞれ「SITE」feature keyと、対応するアミノ酸の省略されていない完全な名前含む「NOTE」qualifierで注釈を付ける必要があります

新しい分子型

D-アミノ酸

```

- <INSDFeature>
  <INSDFeature_key>SITE</INSDFeature_key>
  <INSDFeature_location>1</INSDFeature_location>
  - <INSDFeature_qual>
    - <INSDQualifier id="q4">
      <INSDQualifier_name>NOTE</INSDQualifier_name>
      <INSDQualifier_value>D-alanine</INSDQualifier_value>
    </INSDQualifier>
  </INSDFeature_qual>
</INSDFeature>
- <INSDFeature>
  <INSDFeature_key>SITE</INSDFeature_key>
  <INSDFeature_location>2</INSDFeature_location>
  - <INSDFeature_qual>
    - <INSDQualifier id="q5">
      <INSDQualifier_name>NOTE</INSDQualifier_name>
      <INSDQualifier_value>D-glutamic acid</INSDQualifier_value>
    </INSDQualifier>
  </INSDFeature_qual>
</INSDFeature>
- <INSDFeature>
  <INSDFeature_key>SITE</INSDFeature_key>
  <INSDFeature_location>6</INSDFeature_location>
  - <INSDFeature_qual>
    - <INSDQualifier id="q6">
      <INSDQualifier_name>NOTE</INSDQualifier_name>
      <INSDQualifier_value>D-methionine</INSDQualifier_value>
    </INSDQualifier>
  </INSDFeature_qual>
</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>AEKLG</INSDSeq_sequence>

```

注：同様の例のさらなる考察は
WIPO Standard ST.26, Annex VIのExample 3(a)-1にあります。

新しい分子型

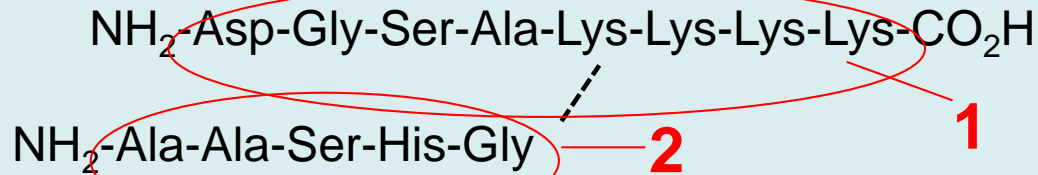
分岐配列

- 分岐核酸配列と分岐アミノ酸配列はST.26の規定の対象です
- 分岐配列の直鎖領域が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を含む場合、その領域を配列表に含める必要があります。(WIPO Standard ST.26, paragraph 7)
- 最小長の要件を満たす分岐配列の各直鎖領域を、独自のSEQ ID 番号を持つ別の配列として含める必要があります
- 構造内の具体的に定義された残基の総数ではなく、個々の直鎖領域における具体的に定義された残基の数を考慮する必要があります

新しい分子型

分岐配列

ある出願で次の配列を持つペプチドが記載されています：



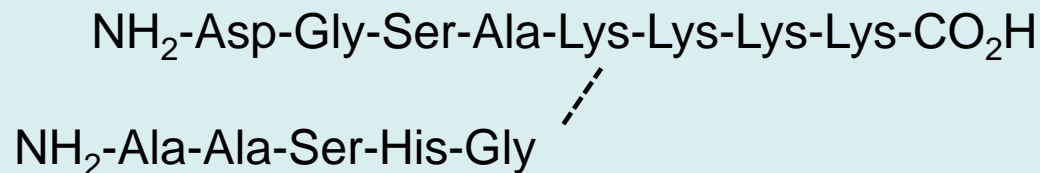
----- はグリシンのカルボキシ末端とリシンの側鎖との間のアミド結合を示します

- ✓ 両方の直鎖領域には4個以上の具体的に定義されたアミノ酸が含まれているため、両方とも配列表に含める必要があります
- ✓ 各直鎖領域を、独自のSEQ ID番号を持つ別の配列として含める必要があります

新しい分子型

分岐配列

ある出願で次の配列を持つペプチドが記載されています：



----- はグリシンのカルボキシ末端とリシンの側鎖との間のアミド結合を示します

- ✓ 両方の直鎖領域には4個以上の具体的に定義されたアミノ酸が含まれているため、両方とも配列表に含める必要があります
- ✓ 各直鎖領域を、独自のSEQ ID番号を持つ別の配列として含める必要があります
- ✓ アミド結合の位置と特徴を示すために、両方の配列に注釈を付ける必要があります

新しい分子型

分岐配列

1

```

- <INSDFeature>
  <INSDFeature_key>SITE</INSDFeature_key>
  <INSDFeature_location>5</INSDFeature_location>
  - <INSDFeature_qual>
    - <INSDQualifier id="q4">
      <INSDQualifier_name>NOTE</INSDQualifier_name>
      <INSDQualifier_value>Lysine side chain is amide bonded to the carboxy
        terminus of the glycine in SEQ ID NO:2</INSDQualifier_value>
    </INSDQualifier>
  </INSDFeature_qual>
</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>DGS AKKKK</INSDSeq_sequence>

```

2

```

- <INSDFeature>
  <INSDFeature_key>SITE</INSDFeature_key>
  <INSDFeature_location>5</INSDFeature_location>
  - <INSDFeature_qual>
    - <INSDQualifier id="q11">
      <INSDQualifier_name>NOTE</INSDQualifier_name>
      <INSDQualifier_value>glycine carboxy terminus is amide bonded to the side
        chain of the lysine in SEQ ID NO:1, position 5</INSDQualifier_value>
    </INSDQualifier>
  </INSDFeature_qual>
</INSDFeature>
</INSDSeq_feature-table>
<INSDSeq_sequence>AASHG</INSDSeq_sequence>

```

注：この例のさらなる考察はWIPO Standard ST.26, Annex VIの
Example 7(b)-3にあります。

変異体配列

変異体配列

- ST.26のparagraph 3(m)では、「**変異体配列**」を

「元の配列に対して1つ以上の差異を含むヌクレオチドまたはアミノ酸配列を指します。この差異には、代替残基（paragraph 15および27参照）、修飾残基（paragraph 3 (g)、3 (h)、16、29 を参照）、欠失、挿入、置換などがあります。paragraph 93～95を参照してください。」と定義しています。

- 変異体配列の開示の仕方によって、配列表でどのように表されなければならないかが決まります。
- ST.26のparagraph 93～95に、変異体をどのように表すべきかが規定されています。

変異体配列

Paragraph 93

Paragraph 93 : それぞれが残基の記載によって開示され、paragraph 7に包含される元の配列およびその配列の変異体は、それぞれが配列表に含まれ、独自の配列識別番号を割り当てられる必要があります。



各変異体が別々に記載される場合、各変異体は独自のSEQ ID番号を持たなければなりません!

変異体配列

Paragraph 93

ある出願に次の多重配列アラインメントを含む図が含まれます：

コンセンサス

Homo sapiens

Pongo abelii

Pan paniscus

Rhinopithecus bieti

Rhinopithecus roxellana

```

LEGneEQFINAakIIRHPkYnrkTlnNDImLIK
LEGNEQFINAAKIIRHPQYDRKTLNNDIMLIK
LEGNEQFINAAKIIRHPQYDRKTVNNDIMLIK
LEGNEQFINAAKIIRHPKYNRITLNNDIMLIK
LEGNEQFINATKIIRHPKYNGNTLNNDIMLIK
LEGNEQFINATQIIRHPKYNGNTLNNDIMLIK
  
```

小文字は、アラインされた配列の中で重要なアミノ酸残基を表します。

- ✓ 6つの記載された配列のそれぞれを、独自のSEQ ID番号を持つ別の配列として配列表に含める必要があります。

注：同様の例のさらなる考察は次にあります
 WIPO Standard ST.26, Annex VI, Example 93-3。

変異体配列

適切な注釈

- 変異体に注釈を付けるための適切なfeature keyは、分子型と変異の特徴によって異なります：

Type of sequence	Feature Key	Qualifier	Use
Nucleic acid	variation	replace or note	Naturally occurring mutations and polymorphisms, e.g., alleles, RFLPs.
Nucleic acid	misc_difference	replace or note	Variability introduced artificially, e.g., by genetic manipulation or by chemical synthesis.
Amino acid	VAR_SEQ	NOTE	Variant produced by alternative splicing, alternative promoter usage, alternative initiation and ribosomal frameshifting.
Amino acid	VARIANT	NOTE	Any type of variant for which VAR_SEQ is not applicable.

WIPO Standard ST.26, paragraph 96

変異体配列

最も制限の厳しいあいまいな記号

- ST.26のparagraph 15および27には、あいまいな記号が必要な場合、「最も制限の厳しい記号を使用しなければならない...」と記載されています

どういう意味か？

以下はその例です：

ヌクレオチド配列において、ある位置が「a または c」であり得る場合、「n」の代わりにあいまいな記号「m」を使用します。

アミノ酸配列において、ある位置が「L または I」であり得る場合、「X」の代わりにあいまいな記号「J」を使用します。

「n」と「X」はデフォルトの値を持っていることに注意してください。したがって、「n」または「X」がデフォルトの値以外に使用される場合は、注釈が必要です。

変異体配列

Paragraph 94

Paragraph 94 : 1つまたは複数の位置に代替残基が記載された単一の配列として開示された変異体配列は、配列表に含める必要があり、単一の配列で表されなければなりません。記載された代替残基は、最も制限の厳しいあいまいな記号で表されます（paragraph 15および27を参照）。



変異体が別々に記載されておらず、元の配列で単に可変残基として表わされている場合、それらは個別のSEQ ID番号を持つ必要はありません！

変異体配列

Paragraph 94

ある出願で次の配列を持つペプチドが開示されています：

Gly-Gly-Gly-[Leu または Ile]-Ala-Thr-[Ser または Thr]

- ✓ 配列表に 1 つの配列として含めることができる
- ✓ 配列は次のように表すのが好ましい：GGGJATX
- ✓ [Leu または Ile] は、最も制限の厳しいあいまいな記号「J」で表されるべきである
- ✓ [Ser または Thr] は、記号「X」とXがセリンまたはスレオニンであることを示すqualifierのnoteを持つfeature key「VARIANT」で表されるべきである

注：この例のさらなる考察は次にあります
WIPO Standard ST.26, Annex VI, Example 94-1。

変異体配列

Paragraph 95

Paragraph 95 : 配列表の元の配列中の欠失、挿入、または置換への参照によってのみ開示される変異体配列は、配列表に含まれていなければなりません。配列表に含まれる場合、そのような変異体は：

- (a) 元の配列の注釈で表すことができ、単一の場所または複数の異なる場所に変異を含み、それらの変異の発生は独立しています；
- (b) 別の配列として表され、独自の配列識別番号を割り当てられなければならない、複数の異なる場所に変異を含み、それらの変異の発生は相互依存しています；
- (c) 別の配列として表され、独自の配列識別番号を割り当てられなければならない、1000残基を超える挿入または置換された配列を含みます（paragraph 86参照）。



「欠失、挿入、または置換への参照」とは、変異体が散文で記述されることを意味します。

変異体配列

Paragraph 95(a)

ある出願に次の開示が含まれています：

ペプチドフラグメント 1 : Gly-Leu-Pro-Xaa-Arg-Ile-Cys

Xaaは任意のアミノ酸

* * *

...別の実施形態では、ペプチドフラグメント 1 はGly-Leu-Pro-Xaa-Arg-Ile-Cys
であり、XaaはVal、Thr、またはAspであり得る.....

* * *

... 別の実施形態では、ペプチドフラグメント 1 はGly-Leu-Pro-Xaa-Arg-Ile-Cys
であり、XaaがValであり得る...

元の配列Gly-Leu-Pro-Xaa-Arg-Ile-Cysは単一の場所に変異が含まれており、それらの変異の発生は独立しています

変異体配列

Paragraph 95(a)

- ✓ 元の配列は配列表に含まれていなければならず、変異体は元の配列の注釈で表すことができます
- ✓ 最も包括的な例は、配列表に含まれる型でなければなりません – この例では、Xaaが「任意のアミノ酸」である型です
- ✓ 配列はGLPXRICと表されなければならず、4位にfeature key「VARIANT」とXaaが「任意のアミノ酸」であることを示すqualifier「note」が必要です
- ✓ 必須ではありませんが、配列表には、次の3つの変異体を別々の配列として含めることを推奨します。

GLPVRIC

GLPTRIC

GLPDRIC

注：この例のさらなる考察はWIPO Standard ST.26, Annex VIの Example 95(a)-1にあります。

変異体配列

Paragraph 95(b)

ある出願で次のコンセンサス配列が記載されています：

aatgn₁cccacgaatgn₂cac

n₁とn₂はa、t、g、またはcであり得る。

いくつかの変異体配列が次のように開示されています：

n₁がaの場合、n₂はt、g、またはc

n₁がtの場合、n₂はa、g、またはc

n₁がgの場合、n₂はt、a、またはc

n₁がcの場合、n₂はt、g、またはa

コンセンサス配列は、複数の異なる場所に変異を含み、それらの変異の発生は相互依存しています

変異体配列

Paragraph 95(b)

- ✓ コンセンサス配列は配列表に含まれていなければならない、変異体は別の配列として表す必要があります
- ✓ 最も包括的な例は、配列表に含まれる型です： n_1 と n_2 は「a、t、g、またはcであり得る」
- ✓ 配列は次のように表されなければなりません：aatgncccacgaatgncac
- ✓ feature keyが存在しない場合、「n」は「a」、「c」、「g」または「t」のいずれかとして解釈されるため、nに対する注釈は不要です（ST.26のparagraph 15を参照）
- ✓ 必須ではありませんが、配列表には、次の4つの変異体を別々の配列として含めることを強く推奨します。

aatgacccacgaatgbcac (b = t, g, またはc)

aatgtvcccacgaatgvcac (v = a, g, またはc)

aatgghcccacgaatghcac (h = t, a, またはc)

aatgdcccacgaatgdcac (b = t, g, またはa)

注：同様の例のさらなる考察は
WIPO Standard ST.26, Annex VIのExample 95(b)にあります。

変異体配列

Paragraph 95(c)

ある出願に次の開示が含まれています：

... -Met-Gly-Leu-Pro-Arg-Xaa-Arg-Ile-Cys-Lys- ...

XaaはGlyまたは配列の挿入

Cys-Tyr-Ile-Lys-Ser-(1000アミノ酸)-Leu-Thr-Pro-Lys

1つの変異体配列に、1000残基を超える挿入または置換された配列が含まれます

変異体配列

Paragraph 95(c)

✓ Xaaが1000以上の残基の挿入である変異体は、独自の配列ID番号を持つ別の配列として配列表に含める必要があります。

✓ XaaがGlyである変異体も、独自の配列ID番号を持つ別の配列として配列表に含めます。

...-MGLPRGRICK-...

変異体配列

Q：次の開示に適用される段落はどれか？

G-L-P-T-R-I-C-[L または I]-A-V-[G または A]

変異体配列

Q : 次の開示に適用される段落はどれか？

G-L-P-T-R-I-C- [L or I] -A-V- [G or A]

A : Paragraph 94 :

「1つまたは複数の位置に代替残基が記載された単一の配列として開示された変異体配列は、配列表に含める必要があり、単一の配列で表されなければなりません。記載された代替残基は、最も制限の厳しいあいまいな記号で表されます。」

変異体配列

Q：次の開示に適用される段落はどれか？

Position	1	2	3	4	5	6	7	8	9
Sequence	A	V	L	T	Y	L	R	G	E
Variant 1									A
Variant 2			P			P			
Variant 3			A	I	G	Y			
Variant 4							-		

表の空白は、変異体に含まれるアミノ酸が「配列」の対応するアミノ酸と同じであることを示し、「-」は「配列」の対応するアミノ酸の欠失を示す。

変異体配列

Q : 次の開示に適用される段落はどれか？

Position	1	2	3	4	5	6	7	8	9
Sequence	A	V	L	T	Y	L	R	G	E
Variant 1									A
Variant 2			P			P			
Variant 3			A	I	G	Y			
Variant 4							-		

表の空白は、変異体に含まれるアミノ酸が「配列」の対応するアミノ酸と同じであることを示し、「-」は「配列」の対応するアミノ酸の欠失を示す。

A : Paragraph 93 :

「それぞれが残基の記載によって開示され、paragraph 7に包含される元の配列およびその配列の変異体は、それぞれが配列表に含まれ、独自の配列識別番号を割り当てられる必要があります。」

変異体配列

Q：次の開示に適用される段落はどれか？

ある図で次の配列が開示されています：

Met-Gly-Ala-Ile-Pro-Asp-Val-Lys-Arg-Ala-Cys-Trp（配列 1）

明細書には、配列 1 に関する次の情報が含まれています：

... 特定の実施形態では、配列 1 の7位のバリンがアラニンで置換されている...

変異体配列

Q : 次の開示に適用される段落はどれか？

ある図で次の配列が開示されています :

Met-Gly-Ala-Ile-Pro-Asp-Val-Lys-Arg-Ala-Cys-Trp (配列 1)

明細書には、配列 1 に関する次の情報が含まれています :

... 特定の実施形態では、配列 1 の7位のバリンがアラニンで置換されている...

A : Paragraph 95(a) : 配列表に元の配列中の欠失、挿入、または置換への参照によってのみ開示される変異体配列は、配。配列表に含まれる場合、そのような変異体は：列表に含まれていなければなりません

(a) 元の配列の注釈で表すことができ、単一の場所または複数の異なる場所に変異を含み、それらの変異の発生は独立しています；

変異体配列

Q：次の開示に適用される段落はどれか？

ある図で次の配列が開示されています：

Met-Gly-Ala-Ile-Pro-Asp-Val-Lys-Arg-Ala-Cys-Trp（配列 1）

明細書には、配列 1 に関する次の情報が含まれています：

...配列 1 の7位のバリンがアラニンに置換されている場合、10位のアラニンはバリンに置換されます....

変異体配列

Q : 次の開示に適用される段落はどれか？

ある図で次の配列が開示されています :

Met-Gly-Ala-Ile-Pro-Asp-Val-Lys-Arg-Ala-Cys-Trp (配列 1)

明細書には、配列 1 に関する次の情報が含まれています :

...配列 1 の7位のバリンがアラニンに置換されている場合、10位のアラニンはバリンに置換されます....

A : Paragraph 95(b) : 配列表に元の配列中の欠失、挿入、または置換への参照によってのみ開示される変異体配列は、配列表に含まれていなければなりません。配列表に含まれる場合、そのような変異体は :

(b) 別の配列として表され、独自の配列識別番号を割り当てられなければならない、複数の異なる場所に変異を含み、それらの変異の発生は相互依存しています ;

質疑応答