

非公式翻訳

注意書き

本書は WIPO 標準 ST.26 の非公式な日本語訳です。

WIPO 国際事務局は本書について責任を負うものではなく、その正確性を検証しているものでもありません。WIPO 国際事務局による検証済みの内容については、以下の WIPO 標準 ST.26 の公式版をご参照ください。

英語 <<https://www.wipo.int/export/sites/www/standards/en/pdf/03-26-01.pdf>>

スペイン語 <<https://www.wipo.int/export/sites/www/standards/es/pdf/03-26-01.pdf>>

フランス語 <<https://www.wipo.int/export/sites/www/standards/fr/pdf/03-26-01.pdf>>

翻訳提供 特許庁

WIPO 標準 ST.26

XML（拡張可能なマークアップ言語）形式で
塩基配列およびアミノ酸配列の配列表を表示するための WIPO 推奨標準

Version 1.5

2021 年 11 月 5 日開催の第 9 回 WIPO 標準委員会(CWS) 会合により採択

WIPO 国際事務局により編修された文書

第 54 回 WIPO 一般総会は、国内、広域及び国際レベルで WIPO 標準 ST.26 に移行する新たなビッグバンの日を、2022 年 7 月 1 日とすることを承認した。
その間、WIPO 標準 ST. 25 は引き続き有効であるものとする。

WIPO 標準は、産業財産権を所管する機関およびその他の関係者に
情報を提供する目的で公開されている。

目次

はじめに	5
用語の定義	5
範囲	7
参考資料	7
配列の表示	8
塩基配列	8
アミノ酸配列	10
特殊な状況の表示	12
XML 形式の配列表の構造	13
ルート要素	14
一般情報部	14
配列データ部	17
feature table	19
feature keys	19
必須の feature keys	19
feature location	19
feature qualifiers	22
必須の feature qualifiers	22
qualifier の要素	22
フリーテキスト	24
コード配列	26
変異体	26

附属書

はじめに

用語の定義

範囲

参考資料

配列の表示

塩基配列

アミノ酸配列

特殊な状況の表示

XML 形式の配列表の構造

ルート要素

一般情報部

配列データ部

feature table

feature keys

必須の feature keys

feature location

feature qualifiers

必須の feature qualifiers

qualifier の要素

フリーテキスト

コード配列

変異体

附属書Ⅱ

附属書Ⅲ

附属書Ⅳ

附属書Ⅴ

附属書Ⅵ

付 録

附属書Ⅶ

WIPO 標準 ST.26

XML（拡張可能なマークアップ言語）形式で

塩基配列表およびアミノ酸配列表を提示するための推奨標準

Version 1.5

2021 年 11 月 5 日開催の第 9 回 WIPO 標準委員会(CWS) 会合で採択

はじめに

1. 本標準は、特許出願において配列表への記載を要する塩基配列およびアミノ酸配列の開示情報。それらの開示情報の表示方法。XML 形式での配列表の文書型の定義(DTD)を定義している。産業財産権を所管する機関は、特許出願の一部として、または特許出願に関連して提出された本標準に準拠型の配列表を受理することを推奨する。
2. 本標準の目的は次のとおりである。
 - (a) 特許出願人が、諸外国および国内または広域的な特許出願手続きの基準を満たす単一の配列表を作成できる様にする。
 - (b) 特許出願人、一般公衆および審査官に利益をもたらす、普及を容易にするため、配列の表示の精度と品質を向上させること。
 - (c) 配列データの検索を容易にすること。および、
 - (d) 配列データを電子的な形式で交換し、コンピューター化されたデータベース上で配列データを利用できるようにすること。

用語の定義

3. 本標準では、次の表現を以下の通り定義する。
 - (a) 「アミノ酸」とは、附属書 I (第 3 節、表 3) に記載されている記号のいずれかを用いて表示できる任意のアミノ酸を意味する。その様なアミノ酸には、特に D-アミノ酸および、修飾を受けた、または合成された側鎖を含むアミノ酸が含まれる。また本標準の第 30 項に従って feature table に修飾を受けたものとして更に説明がない限り、アミノ酸とは非修飾の L-アミノ酸として解釈される。また本標準ではペプチド核酸(PNA)の残基はアミノ酸であるとみなされず、第 3 項(g)(i)(2)で定められている通りヌクレオチドであるとみなされる。
 - (b) 「統制語彙」とは本標準で使用されている用語であり、配列の feature（配列の特徴）に関して記載する際に用いられなければならない用語である。すなわち、附属書 I で定められている通り、対象の領域や部位の注釈に用いられる用語である。
 - (c) 「残基の数え上げ」とは、特許出願において、各残基配列を順に数え上げて開示することを意味し、ここでは以下のような。
 - (i) 残基が名称、略語、記号または構造で示されている(例. HHHHHHQ また HisHisHisHisHisHisGln)。または、

- (ii) 複数の残基が簡略な式で表されている（例. His₆Gln）。
- (d) 「意図的にスキップされた配列」は空の配列としても知られており、出願開示内容との一貫性を保つために、配列表上の配列の番号付けを維持するための記入子を指し、例えば、開示内容と配列表の両方で配列の番号の振り直しを避けるため、開示内容から配列が削除される場合等に使用される。
- (e) 「修飾アミノ酸」とは、第 3 項(a)で説明されている、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-セレンシステイン、L-トレオニン、L-トリプトファン、L-チロシンまたは L-バリン以外のアミノ酸を意味する。
- (f) 「修飾塩基」とは、デオキシアデノシン 3'-リン酸塩、デオキシグアノシン 3'-リン酸塩、デオキシチジン 3'-リン酸塩、デオキシチミジン 3'-リン酸塩、アデノシン 3'-リン酸塩、グアノシン 3'-リン酸塩、シチジン 3'-リン酸塩またはウリジン 3'-リン酸塩以外の第 3(g)項に記載されたいずれかのヌクレオチドのことである。
- (g) 「ヌクレオチド」とは、ヌクレオチドまたはヌクレオチド類似体に関する附属書 I (第 1 節、表 1)で定められている記号のいずれかを用いて表示できる、任意のヌクレオチドまたはヌクレオチド類似体のことである。
- (i) 骨格部分は以下のいずれかから選択する。
- (1) 2'デオキシリボース 5'リン酸（デオキシリボヌクレオチドの主鎖の一部）または、リボース 5' リン酸（リボヌクレオチドの主鎖の一部）。または、
- (2) 核酸類似体の主鎖を形成するときに、2'デオキシリボース 5'リン酸またはリボース 5'リン酸の主鎖を含む核酸における核酸塩基の配列を模倣する核酸の配列をもたらす、2'デオキシリボース 5'リン酸またはリボース 5'リン酸の類似体であって、ここで、核酸類似体は相補的な核酸と塩基対をなすことができるものである。骨格部分の例として、ペプチド核酸のようなアミノ酸、グリコール核酸のようなグリコール分子、スレオース核酸のようなスレオフラノシル糖分子、モルホリノのようなモルホリン環およびホスホロジアミデート基、およびシクロヘキセニル核酸のようなシクロヘキセニル分子などがある。
- また、
- (ii) 骨格部分は以下のいずれかとなる。
- (1) 修飾を受けたまたは合成されたプリンまたはピリミジン核酸塩基を含む、核酸塩基と結合したものの。
- または、
- (2) ヌクレオチドが「AP 部位」または「脱塩基部位」と呼ばれる塩基配列の一部である場合、プリンまたはピリミジン核酸塩基を欠いているもの。
- (h) 「残基」とは、配列内の任意の個々のヌクレオチドまたはアミノ酸、またはこれらの類似体を意味する。
- (i) 「配列番号」とは、配列表の各配列に割り当てられた、その配列固有の番号（整数）を意味する。
- (j) 「配列表」とは、出願時に提出される特許出願明細書の一部、または出願後に提出される書類であって、本標準に規定されている開示されたヌクレオチドおよび/またはアミノ酸の配列を、詳細な説明とともに包含するものを意味する。
- (k) 「具体的に定義された」とは、記号「n」で表されるヌクレオチド以外の任意のヌクレオチドおよび記号「X」で表されるアミノ酸以外の任意のアミノ酸であって、附属書 I に記載されているものを意味する（それぞれ、第 1 項表 1 および第 3 項表 3 を参照のこと）。
- (l) 「不明な」ヌクレオチドまたはアミノ酸とは、単一のヌクレオチドあるいはアミノ酸が存在するが、その同一性が不明であるか、または開示されていないものをいう。
- (m) 「変異体配列」とは、一次配列に関して 1 つまたは複数の差異を含む塩基配列またはアミノ酸配列を意味す

る。これらの違いには、代替残基（第 15 項および第 27 項を参照）、改変残基（第 3 項 (g)、第 3 項 (h)、第 16 項、および第 29 項を参照）、欠失、挿入、および置換が含まれる場合がある。第 93 項から第 95 項参照のこと。

(n)「フリーテキスト」は、特定の **qualifier** 値の形式の一種であり、説明文のフレーズまたはその他の指定された形式（附属書 I に示されている）の形式で表示される。第 85 項参照のこと。

(o)「言語依存のフリーテキスト」とは、特定の **qualifier** のフリーテキスト値を意味する。これには、国内的、広域的、または国際的な手順の翻訳が必要になる場合がある。第 87 項参照のこと。

4. 本標準では次の言葉を以下の通り定義する。

- (a) 法助動詞の「may」は、可能または許容を意味するが、必須ではない。
- (b) 法助動詞の「must」は本標準の定める要件であり、この要件を満たさない場合は不適合となる。
- (c) 法助動詞の「must not」は、本標準の禁止事項を表している。
- (d) 法助動詞の「should」は、強い推奨を意味するが、必須ではない。
- (e) 法助動詞の「should not」は、強い制止を意味するが、禁止ではない。

範囲

5. 本標準は、特許出願において開示されている配列のヌクレオチドおよびアミノ酸の配列表を提示するための要件を定めるものである。

6. 本標準に準拠する配列表（以降、「配列表」という）は一般情報部と、配列データ部とを含んでいる。配列表は、附属書 II に示されている文書型定義(DTD)を用いた XML 形式の単一ファイルとして提示しなければならない。一般情報部に含まれる書誌情報の目的は、配列表と、その配列表が提出された特許出願とを紐づけることのみにある。配列データ部は、1 つまたは複数の配列データ要素で構成され、それぞれが 1 の配列に関する情報を含む。配列データ要素は、国際塩基配列データベースコラボレーション(INSDC)および Uniprot（アミノ酸配列とその機能情報を掲載しているデータベース）仕様に基づく様々な **feature keys** とその後続の **qualifiers** を含むものである。

7. 本標準では、配列表への記載を要する配列とは、その残基の数え上げによって出願のどこにでも開示され、次のように表示できる配列である。

(a) 10 個以上の具体的に定義されたヌクレオチドを含む非分岐配列または分岐配列の線形領域で、次の方法で隣接するヌクレオチドと結合する。

- (i) 3'位から 5'位の間（または 5'位から 3'位の間）のホスホジエステル結合。
または
- (ii) 天然に存在する核酸中の核酸塩基の配置を模倣する隣接する核酸塩基の配置をもたらす化学結合。

(b) 4 個以上の具体的に定義されたアミノ酸を含む非分岐配列または分岐配列の線形領域であって、その 4 つ以上の具体的に定義されたアミノ酸が結合して単一のペプチド主鎖を形成する（すなわち、隣接するアミノ酸がペプチド結合により結合する）。

8. 配列表には、個別の配列番号が付与された配列として、具体的に定義された 10 個未満のヌクレオチドまたは、具体的に定義された 4 個未満のアミノ酸の配列を記載してはならない。

参考資料

9. 以下の標準および利用可能な情報に関する資料は、本標準と関連性がある。

国際塩基配列データベースコラボレーション(INSDC)	http://www.insdc.org/
国際標準化機構(ISO)言語コード 639-1: 2002 年公表	言語名を表示するためのコード – パート 1 : アルファベット 2 文字のコード
Unitprot (蛋白質に関する共通情報)コンソーシアム	http://www.uniprot.org/
W3C XML 1.0	http://www.w3.org/

WIPO Standard ST.2	グレゴリオ暦を使用してカレンダーの日付を指定する方法
WIPO Standard ST.3	国、その他の団体および政府間組織を 2 文字で表すための推奨規格
WIPO Standard ST.16	異なる種類の特許文献を識別するための推奨標準コード
WIPO Standard ST.25	特許出願におけるヌクレオチドおよびアミノ酸の配列表の表示に関する標準

配列の表示

- 1 0. より長い配列の領域と同一の配列を含め、第 7 項に記載の各配列には、個別の配列番号を割り振らなければならない。配列番号は「1」から始まり、以降、整数の連番で増加する必要がある。配列識別番号は存在するがその番号に該当する配列が存在しない場合、すなわち意図的にスキップされた配列では、その配列の代わりに「000」を用いなければならない（第 58 項参照）。配列の総数は、配列が続くか、「000」が続くかにかかわらず、配列番号の総数と等しくなければならない。

塩基配列

- 1 1. 塩基配列は、左から右への 5'位から 3'位方向、または 5'位から 3'位方向を模した左から右への一本鎖によってのみ表されなければならない。5'位および 3'位または他の同様の指定を配列に記載してはならない。両鎖の残基の数え上げによって開示された二本鎖のヌクレオチド配列は次のように表示しなければならない。
- (a) 2 つの別々の鎖が互いに完全に相補的である場合、単一の配列、または 2 つの別々の配列としてそれぞれに個別の配列番号が割り当てられる。または、
- (b) 2 つの鎖が互いに完全には相補的でない場合、2 つの別々の配列であり、それぞれに個別の配列番号が割り当てられる。
- 1 2. 本標準では、配列に示される最初のヌクレオチドは残基位置番号 1 である。塩基配列が環状構造である場合、出願人は残基位置番号 1 のヌクレオチドを選択しなければならない。配列番号の割り振りは、5'末端から 3'末端の方向、または 5'末端から 3'末端の方向を模した方向に、配列全体に連番で割り振る。最後の残基位置番号は、配列内のヌクレオチド数と等しくなければならない。
- 1 3. 配列内のすべてのヌクレオチドは附属書 I（第 1 節、表 1 を参照のこと）で定められている記号を用いて表示しなければならない。その際は、アルファベットの小文字のみを用いて表示しなければならない。ヌクレオチドの表示に用いる記号は、1 つの残基のみを示す。
- 1 4. 記号「t」は、DNA ではチミン、RNA ではウラシルと解釈される。DNA のウラシルや RNA のチミンは、修飾塩基とみなされ、第 19 項で規定された feature table でさらに説明しなければならない。
- 1 5. 曖昧さ記号（2 つ以上の代替ヌクレオチドを表す）が適切な場合は、附属書 I（第 1 節、表 1）に記載されているように、最も限定的な記号を使用すべきである。例えば、特定の位置のヌクレオチドが「a」または「g」である可能性がある場合は、「n」ではなく「r」を使用すべきである。記号「n」は、feature table でさらに説明する場合を除いて、「a」、「c」、「g」または「t/u」のいずれかの 1 つであると解釈される。記号「n」は、ヌクレオチド以外のものを表すために用いてはならない。単一の修飾を受けた、または「不明な」ヌクレオチドは、第 16 項、17 項、21 項、または第 93 項から第 96 項に記載されているように、feature table におけるさらなる説明とともに記号「n」で表すことができる。配列変化、すなわち代替、欠失、挿入または置換の表現については、第 93 項から第 100 項を参照のこと。
- 1 6. 修飾塩基は、可能な限り対応する非修飾塩基のヌクレオチド、すなわち「a」、「c」、「g」または「t」として配列内で表現すべきである。配列内の修飾塩基のうち、附属書 I（第 1 節、表 1 参照）の他の記号で表すことができないもの、すなわち天然由来でないヌクレオチド等の「他の」ヌクレオチドは、記号「n」で表さなければならない。記号「n」は 1 つの残基のみを示す。
- 1 7. 修飾塩基に関しては、feature key「modified_base」およびその必須 qualifier「mod_base」を使用して、qualifier の値として附属書 I（第 2 節、表 2 参照）の略語の 1 つと組み合わせて、feature table（第 60 項以降参照）に詳しく説明しなければならない。略語が「OTHER」である場合は、その修飾塩基の完全で、省略されていない名称を、「note」qualifier の値として指定しなければならない。代替の修飾塩基のリストについては、qualifier 値「OTHER」をさらに「note」qualifier と組み合わせて使用しても良い（第 97 および 98 項参照）。前述の附属書 I（第 2 節、表 2 参照）に記載されている略語（または完全名称）は、配列自体で使用してはならない。

- 1 8. 同じ骨格部分を共有する連続した修飾塩基の 1 つまたは複数の領域を含む塩基配列（第 3 項(g)(i)(2)参照）は、第 17 項の規定のとおり、さらに詳しく feature table で説明しなければならない。また第 22 項で規定されているように、領域内のすべての修飾塩基は、単一の INSDFeature 要素でまとめて記載しても良い。領域内のすべての修飾塩基を包含する最も限定的で省略されていない化学名または領域内のすべてのヌクレオチドの化学名のリストは、「note」qualifier の値として提供しなければならない。例えば、「a」、「c」、「g」または「t」の核酸塩基を含むグリコール核酸配列は、「note」qualifier に、「2,3-ジヒドロキシプロピルヌクレオシド」として記載し得る。あるいは、同じ配列を、「note」qualifier に、「2,3-ジヒドロキシプロピルアデニン、2,3-ジヒドロキシプロピルチミン、2,3-ジヒドロキシプロピルグアニン、2,3-ジヒドロキシプロピルシトシン」として記載し得る。領域内の個々の修飾塩基が追加の修飾を含む場合、その修飾塩基も、第 17 項に規定されているように、feature table に詳しく説明しなければならない。
- 1 9. DNA のウラシルや RNA のチミンは修飾塩基であるとみなされる。配列上では、「t」と表記されなければならない。さらに、feature table では、feature key「modified_base」、qualifier の値として qualifier「mod_base」に「OTHER」を組み合わせて、および、qualifier 値として qualifier「note」にそれぞれ「uracil」または「thymine」を組み合わせて表記されなければならない。
- 2 0. 次の例は、上記第 16 項から第 18 項による修飾塩基の表示例を示す。

例 1：附属書 I の略語を用いた修飾塩基（第 2 節、表 2 参照）

```
<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>
```

例 2：附属書 I の「OTHER」を用いた修飾塩基（第 2 節、表 2 参照）

```
<INSDFeature>
  <INSDFeature_key>modified_base</INSDFeature_key>
  <INSDFeature_location>4</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>
```

例 3：第 3 項(g)(i)(2)に包含される修飾塩基からなる塩基配列であって、さらに修飾を受けた 2 つの個別のヌクレオチドを有するもの

```
<INSDFeature>
  <INSDFeature_key>modified_base</INSDFeature_key>
  <INSDFeature_location>1..954</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</INSDQualifier_value>
    </INSDQualifier>
  </INSDFeature_qual>
</INSDFeature>
<INSDFeature>
  <INSDFeature_key>modified_base</INSDFeature_key>
  <INSDFeature_location>439</INSDFeature_location>
  <INSDFeature_qual>
    <INSDQualifier>
      <INSDQualifier_name>mod_base</INSDQualifier_name>
      <INSDQualifier_value>i</INSDQualifier_value>
    </INSDQualifier>
  </INSDFeature_qual>
</INSDFeature>
<INSDFeature>
  <INSDFeature_key>modified_base</INSDFeature_key>
  <INSDFeature_location>684</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>

```

- 2 1. 「不明な」ヌクレオチドはすべて、配列内の記号「n」で表さなければならない。また「不明な」ヌクレオチドは feature key「unsure」を使用して、feature table（第 60 項以降参照）でさらに詳しく説明すべきである。記号「n」は、1 つの残基のみを示す。
- 2 2. 同じ記載が適用される既知数の連続した「a」、「c」、「g」、「t」または「n」残基を含む領域は、要素 INSDFeature_location に存在位置記述子として、「x...y」という構文を持つ単一の要素 INSDFeature element を用いてまとめて記載し得る（第 64 項から第 71 項参照）。配列変化、すなわち、代替、欠失、挿入または置換の表現については、第 93 項から第 100 項を参照のこと。
- 2 3. 次の例は、上記第 22 項に従い、同一の記載が適用される修飾塩基の領域の表現を示している。

```

<INSDFeature>
  <INSDFeature_key>modified_base</INSDFeature_key>
  <INSDFeature_location>358..485</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>isoguanine</INSDQualifier_value>
    </INSDQualifier>
  </INSDFeature_qual>
</INSDFeature>

```

アミノ酸配列

24. アミノ酸配列のアミノ酸は、左から右に向かって、アミノ基からカルボキシ基の方向で表されなければならない。アミノ基およびカルボキシ基は配列内に表示してはならない。
25. 本標準では、配列の最初のアミノ酸は残基番号 1 であり、成熟蛋白質に先行するアミノ酸、例えば、プレ配列、プロ配列、プレプロ配列、シグナル配列なども含まれる。アミノ酸配列が環状で、その環がペプチド結合で結合されたアミノ酸残基のみからなる場合、すなわち、配列にアミノ末端およびカルボキシ末端が無い場合、出願人は残基位置 1 のアミノ酸を選択しなければならない。番号の割り振りは、配列全体を通して、アミノ基からカルボキシ基の方向に連続する。
26. 配列内のすべてのアミノ酸は、附属書 I（第 3 節、表 3 参照）に記載されている記号を用いて表示しなければならない。アルファベットの大文字のみを使用しなければならない。アミノ酸の表示に用いる記号は、1 つの残基のみを示す。
27. 曖昧さ記号（代替となる 2 以上のアミノ酸を表す）が適切である場合には、附属書 I（第 3 章、表 3）に記載されているように、最も限定的な記号を使用すべきである。例えば、特定の位置のアミノ酸がアスパラギン酸またはアスパラギンである可能性がある場合には、「X」ではなく「B」という記号を使用すべきである。記号「X」は、feature table の詳細な情報とともに使用される場合を除き、「A」、「R」、「N」、「D」、「C」、「Q」、「E」、「G」、「H」、「I」、「L」、「K」、「M」、「F」、「P」、「O」、「S」、「U」、「T」、「W」、「Y」、または「V」のいずれか 1 つと解釈される。記号「X」は、アミノ酸以外のものを表すために使用してはならない。単一の修飾アミノ酸または「不明な」アミノ酸は、例えば、第 29 項、第 30 項、第 32 項、または第 92 項から第 98 項で提供されるような feature table に詳しい説明とともに、記号「X」で表すことができる。配列変化、すなわち、代替、欠失、挿入または置換の表現については、第 93 項から第 100 項を参照のこと。
28. 例えば、「Ter」、アスタリスク「*」、ピリオド「.」または空白「 」で表される内部終止記号で区切られた開示されたアミノ酸配列は、少なくとも 4 つの具体的に定義されたアミノ酸を含み、かつ第 7 項に包含される各アミノ酸配列について、個別の配列として含めなければならない。このような個別の配列には、それぞれ独自の配列番号を割り当てなければならない。終止記号および空白は、配列表の配列に含めてはならない(第 57 項参照)。
29. D-アミノ酸を含む修飾アミノ酸は、可能な限り、対応する非修飾アミノ酸として配列に表示すべきである。配列の修飾アミノ酸のうち、附属書 I（第 3 節、表 3 参照）の他の記号で表すことができないもの、すなわち、「その他」のアミノ酸は、「X」で表示しなければならない。記号「X」は、1 つの残基のみを示す。
30. 修飾アミノ酸は、feature table にさらに記述しなければならない（第 60 項以降を参照のこと）。必要に応じて、feature key「CARBOHYD」または「LIPID」は、「note」qualifier と共に使用すべきである。その他の翻訳後の修飾アミノ酸については、「note」qualifier と共に feature key「MOD_RES」を使用すべきであり、それ以外の場合は、feature key「SITE」を「note」qualifier とともに使用する必要がある。「note」qualifier の値は、附属書 I に記載されている略語（第 4 節、表 4 参照）または修飾アミノ酸の省略されていない完全な名称のいずれかでなければならない。上記の表 4 に記載されている略語や、省略されていない完全な名称は、配列自体に使用してはならない。
31. 次の例は、上記第 30 項による修飾アミノ酸の表示である。

例 1：翻訳後修飾アミノ酸

```
<INSDFeature>
  <INSDFeature_key>MOD_RES</INSDFeature_key>
  <INSDFeature_location>3</INSDFeature_location>
  <INSDFeature_qual>
    <INSDQualifier>
      <INSDQualifier_name>note</INSDQualifier_name>
      <INSDQualifier_value>3Hyp</INSDQualifier_value>
    </INSDQualifier>
  </INSDFeature_qual>
</INSDFeature>
```

例 2：翻訳後非修飾アミノ酸

```
<INSDFeature>
  <INSDFeature_key>SITE</INSDFeature_key>
  <INSDFeature_location>3</INSDFeature_location>
  <INSDFeature_qual>
    <INSDQualifier>
      <INSDQualifier_name>note</INSDQualifier_name>
      <INSDQualifier_value>Orn</INSDQualifier_value>
    </INSDQualifier>
  </INSDFeature_qual>
</INSDFeature>
```

例 3 : D-アミノ酸

```
<INSDFeature>
  <INSDFeature_key>SITE</INSDFeature_key>
  <INSDFeature_location>9</INSDFeature_location>
  <INSDFeature_qual>
    <INSDQualifier>
      <INSDQualifier_name>note</INSDQualifier_name>
      <INSDQualifier_value>D-Arginine</INSDQualifier_value>
    </INSDQualifier>
  </INSDFeature_qual>
</INSDFeature>
```

3 2. 「不明な」アミノ酸は、配列に記号「X」を用いて表示しなければならない。「X」と指定された「不明な」アミノ酸は、feature key「UNSURE」と、任意で「note」qualifier を用いて、feature table (第 60 項以降参照) にさらに記述しなければならない。記号「X」は、1 つの残基のみを示す。

3 3. 次の例は、上記第 32 項による「不明な」アミノ酸の表示である。

```
<INSDFeature>
  <INSDFeature_key>UNSURE</INSDFeature_key>
  <INSDFeature_location>3</INSDFeature_location>
  <INSDFeature_qual>
    <INSDQualifier>
      <INSDQualifier_name>note</INSDQualifier_name>
      <INSDQualifier_value>A or V</INSDQualifier_value>
    </INSDQualifier>
  </INSDFeature_qual>
</INSDFeature>
```

3 4. 同じ記述が適用される既知数の連続した「X」残基を含む領域は、要素 INSDFeature_location の□ケーション記述子として「x.y」という構文を用いてまとめて記載しても良い(第 64 項から 70 項参照)。配列変化、すなわち、代替、欠失、挿入、または置換の表現については、第 93 項から 100 項を参照のこと。

特殊な状況の表示

3 5. 残基数の数え上げによって開示された配列で、より長い配列の 1 つまたは複数の非連続セグメント、または異なる配列のセグメントから単一の連続配列として構成されているものは、配列表に記載しなければならず、個別の配列番号を割り当てなければならない。

3 6. 具体的に定義された残基の領域が、連続した「n」または「X」残基の 1 つまたは複数の領域によって分離されている配列 (それぞれ第 15 項および第 27 項参照) で、各領域の「n」または「X」残基の正確な数が開示されているものは、単一の配列として配列表に記載しなければならず、個別の配列番号を割り当てなければならない。

3 7. 具体的に定義された残基の領域が、不明のまたは未公表の数の残基の 1 つまたは複数のギャップによって分離されている配列は、単一の配列として配列表に記載してはならない。第 7 項に記載されている特異的に定義された残基の各領域は、個別の配列として配列表に記載しなければならず、個別の識別番号が割り当てられなければならない。

XML 形式の配列表の構造

38. 上記第6項に従い、本標準に準拠型の配列表ファイルのXMLインスタンスは次のとおりである。

- (a) 一般情報部：配列表の対象となる特許出願に関する情報を含み、および、
- (b) 配列データ部：1つまたは複数の配列データ要素を包み、各要素には1つの配列に関する情報を含む。

配列表の例は、附属書IIIに記載されている。

39. 配列表は、附属書II「配列表の文書型定義(DTD)」に示すDTDを用いて、XML (ver1.0) で提示しなければならない。

- (a) XMLインスタンスの1行目には、XML宣言を記載しなければならない。

```
<?xml version="1.0" encoding="UTF-8"?>
```

- (b) XMLインスタンスの2行目には、文書型(DOCTYPE)宣言が含まれていなければならない。

```
<!DOCTYPE ST26SequenceListing PUBLIC "-//WIPO//DTD Sequence Listing 1.3 //EN"
"ST26SequenceListing_V1_3.dtd">
```

40. 電子配列表全体を1つのファイルに含めなければならない。ファイルはUnicode UTF-8を使用してエンコードしなければならないが、次の制限がある。

- (a) 一般情報部の要素であるApplicantName、InventorName、InventionTitle、および配列データ部のNonEnglishQualifier_valueに含まれる情報は、Unicode Control コードポイント 0000-001F および 007F-009F を除き、XML 1.0 仕様で示される有効な Unicode 文字で構成し得る。予約文字「」、「&」、「'」、「<」、「>」（それぞれ、Unicode コードポイント 0022、0026、0027、003C、003E）は、第41項に記載されているように置き換えなければならない。そして、
- (b) 一般情報部の他のすべての要素および属性、ならびに配列データ部の他のすべての要素および属性に含まれる情報は、Unicode 基本ラテンコード表の印刷可能な文字（スペース文字を含む）で構成されていなければならない（すなわち、Unicode コードポイント 0020 から 007E までに限定される- 附属書IV参照）。予約文字「」、「&」、「'」、「<」、「>」（それぞれ Unicode コードポイント 0022、0026、0027、003C、003E）は、第41項に記載されているように置き換えなければならない。

41. 配列表のXMLインスタンスでは、数値文字参照¹を使用してはならず、要素の属性または内容の値で使用する場合は、次の予約文字を対応する定義済み実体に置き換えなければならない。

予約文字	定義済み実体
<	<
>	>
&	&
"	"
'	'

許可される文字参照は、この項で規定されている定義済み文字実体のみである。例として、第71項を参照のこと。

42. すべての必須要素にデータを入力しなければならない（意図的にスキップされた配列について第58項で規定されている場合を除く）。内容が利用できない任意要素は、XMLインスタンスに表示すべきではない（配列での「replace」qualifierの値の削除表現について第97項で規定されている場合を除く）。

¹ 数値文字参照は、そのユニバーサル文字セット/ユニコードコードポイントによって文字を参照し、「& # nnnn;」または「& # xhhhh;」の形式を使用する。ここで、「nnnn」は10進形式のコードポイントであり、「hhhh」は16進形式のコードポイントである。

ルート要素

4 3. 本標準に準拠型の XML インスタンスのルート要素は、次の属性を持つ要素 ST26SequenceListing である。

属性	記載情報	必須/任意
dtdVersion	「V#.#」の形式でこのファイルを作成するために使用する DTD のバージョン、e.g., 「V1_3」。	必須
fileName	配列表ファイルの名称	任意
softwareName	ファイルの生成に用いたソフトウェア名	任意
softwareVersion	ファイルの生成に用いたソフトウェアのバージョン	任意
productionDate	配列表ファイルの作成日(ひな型は、「CCYY-MM-DD」)	任意
originalFreeTextLanguageCode	言語に依存するフリーテキスト qualifier が作成された単一の原語の言語コード (第 9 項の ISO 639-1:2002 への参照文を参照)	任意
nonEnglishFreeTextLanguageCode	要素 NonEnglishQualifier_value の言語コード (第 9 項の ISO 639-1:2002 への参照文を参照)	配列表に要素 NonEnglishQualifier_value が存在する場合は必須

4 4. 次の例は、上記第 43 項の XML インスタンスのルート要素 T26SequenceListing およびその属性を示している。

```
<ST26SequenceListing dtdVersion="V1_3" fileName="US11_405455_SEQ1.xml"
softwareName="WIPO Sequence" softwareVersion="1.0" productionDate="2022-05-10"
originalFreeTextLanguageCode="de" nonEnglishFreeTextLanguageCode="fr">
{...}*
</ST26SequenceListing>
```

*{...}は、この例に含まれていない一般情報部と配列データ部を示している。

一般情報部

4 5. 一般情報部の要素は、次のように特許出願情報に関連している。

要素	記載情報	必須/任意
ApplicationIdentification	特許出願番号 (配列表はこの特許出願番号で提出される)	出願番号の割り当て後、配列表を提出する場合は必須
The ApplicationIdentification (以下で構成される)		
IPOfficeCode	特許出願した知財庁の ST.3 コード	必須
ApplicationNumberText	特許出願した知財庁が割り当てた特許出願番号(例. PCT/IB2013/099999)	必須
FilingDate	配列表が提出された特許出願の出願日(ST.2 のひな型「CCYY-MM-DD」、すなわち暦年(4桁)、暦月(2桁)および暦日の暦日(2桁)、例. 2015-01-31)	出願日の確定後は、配列表を提出する場合は必須
ApplicantFileReference	第 40 項(b)に記載のある文字で入力された、出願人が特定の出願を識別するために割り当てる個別の識別番号	出願番号の割り当て前の時点で配列表を提出する場合は必須、それ以外は任意

EarliestPriorityApplicationIdentification	最先の優先権出願の識別情報 (IPOfficeCode、ApplicationNumberText および FilingDate も含む。上記 ApplicationIdentification 参照)	優先権が主張されている場合は必須
ApplicantName	第 40 項(a)に記載のある文字で入力された最初に挙げられた出願人の名前。この要素には、第 47 項に規定する必須属性 languageCode が含まれる。	必須
ApplicantNameLatin	ApplicantName が第 40 項(b)に記載のある文字以外で入力されている場合、最初に挙げられた出願人名の翻訳または翻字は第 40 項(b)に定める文字でも入力しなければならない。	出願人名にラテン文字以外の文字が含まれている場合は必須
InventorName	第 40 項(a)に記載のある文字で入力された最初に挙げられた発明者名。この要素には、第 47 項にきさいのある必須属性 languageCode が含まれる。	任意
InventorNameLatin	InventorName が第 40 項(b)に記載のある文字以外で入力されている場合、最初に挙げられた発明者の翻訳または翻字は第 40 項(b)に定める文字で入力しても良い。	任意
InventionTitle	出願言語で、第 40 項(a)に記載のある文字で入力された発明の名称。追加言語への発明の名称の翻訳は、要素 Invention Title を使用して、第 40 項(a)に記載のある文字で入力しても良い。この要素には、第 48 項に記載のある必須属性 languageCode を含む。発明の名称は 2 文字から 7 文字とされるべきである。	出願言語で記入する場合は必須、追加言語で記入する場合は任意
SequenceTotalQuantity	意図的にスキップされた配列 (空の配列とも呼ばれる) を含む、配列表の全配列の総数 (第 10 項参照)。	必須

4 6. 次の例は、上記の第 45 項に準拠型の配列表の一般情報部の表示を示す。

例 1 : 出願識別情報および出願日が付与される前に提出された配列表

```
<?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="Invention_SEQ1.xml" softwareName="WIPO Sequence"
softwareVersion="1.0" productionDate="2022-05-10"
originalFreeTextLanguageCode="en" nonEnglishFreeTextLanguageCode="ja">
  <ApplicantFileReference>AB123</ApplicantFileReference>
  <EarliestPriorityApplicationIdentification>
    <IPOfficeCode>IB</IPOfficeCode>
    <ApplicationNumberText>PCT/IB2013/099999</ApplicationNumberText>
    <FilingDate>2014-07-10</FilingDate>
  </EarliestPriorityApplicationIdentification>
  <ApplicantName languageCode="en">GENOS Co., Inc.</ApplicantName>
```

```

<InventorName languageCode="en">Keiko Nakamura</InventorName>
<InventionTitle languageCode="en">SIGNAL RECOGNITION PARTICLE RNA AND
PROTEINS</InventionTitle>
<SequenceTotalQuantity>9</SequenceTotalQuantity>
<SequenceData sequenceIDNumber="1"> {...}*</SequenceData>
<SequenceData sequenceIDNumber="2"> {...}</SequenceData>
<SequenceData sequenceIDNumber="3"> {...}</SequenceData>
<SequenceData sequenceIDNumber="4"> {...}</SequenceData>
<SequenceData sequenceIDNumber="5"> {...}</SequenceData>
<SequenceData sequenceIDNumber="6"> {...}</SequenceData>
<SequenceData sequenceIDNumber="7"> {...}</SequenceData>
<SequenceData sequenceIDNumber="8"> {...}</SequenceData>
<SequenceData sequenceIDNumber="9"> {...}</SequenceData>
</ST26SequenceListing>

```

*{...} は、この例に含まれていない各配列の関連情報を表している。

例 2：出願識別番号および出願日が付与された後に提出された配列表

```

<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE ST26SequenceListing PUBLIC "-//WIPO//DTD Sequence Listing 1.3//EN"
"ST26SequenceListing_V1_3.dtd">
<ST26SequenceListing dtdVersion="1_3" fileName="Invention_SEQ1.xml" softwareName="WIPO Sequence"
softwareVersion="1.0" productionDate="2022-05-10"
originalFreeTextLanguageCode="en" nonEnglishFreeTextLanguageCode="ja">
  <ApplicationIdentification>
    <IPOfficeCode>US</IPOfficeCode>
    <ApplicationNumberText>14/999,999</ApplicationNumberText>
    <FilingDate> 2015-01-05</FilingDate>
  </ApplicationIdentification>
  <ApplicantFileReference>AB123</ApplicantFileReference>
  <EarliestPriorityApplicationIdentification>
    <IPOfficeCode>IB</IPOfficeCode>
    <ApplicationNumberText>PCT/IB2014/099999</ApplicationNumberText>
    <FilingDate>2014-07-10</FilingDate>
  </EarliestPriorityApplicationIdentification>
  <ApplicantName languageCode="en">GENOS Co., Inc.</ApplicantName>
  <InventorName languageCode="en">Keiko Nakamura</InventorName>
  <InventionTitle languageCode="en">SIGNAL RECOGNITION PARTICLE RNA AND
PROTEINS</InventionTitle>
  <SequenceTotalQuantity>9</SequenceTotalQuantity>
  <SequenceData sequenceIDNumber="1"> {...}*</SequenceData>
  <SequenceData sequenceIDNumber="2"> {...}</SequenceData>
  <SequenceData sequenceIDNumber="3"> {...}</SequenceData>
  <SequenceData sequenceIDNumber="4"> {...}</SequenceData>
  <SequenceData sequenceIDNumber="5"> {...}</SequenceData>
  <SequenceData sequenceIDNumber="6"> {...}</SequenceData>
  <SequenceData sequenceIDNumber="7"> {...}</SequenceData>
  <SequenceData sequenceIDNumber="8"> {...}</SequenceData>
  <SequenceData sequenceIDNumber="9"> {...}</SequenceData>
</ST26SequenceListing>

```

*{...} は、この例に含まれていない各配列の関連情報を表している。

- 4 7. 出願人名および、任意に発明者名は、通常、出願言語で参照されるため、要素 ApplicantName および InventorName でそれぞれ示さなければならない。適切な言語コード（第 9 項の「国際標準化機構(ISO) 言語コード 639-1:2002 年公表」の参照を参照）は、各要素の属性 languageCode で示さなければならない。示された出願人名に、第 40 項(b)に記載のあるラテン文字以外の文字を含む場合、出願人名の翻字または翻訳は要素 ApplicantNameLatin で、ラテン文字でも示さなければならない。示された発明者名がラテン文字以外の文字を含む場合、発明者名の翻字または翻訳は、要素 InventorNameLatin で、ラテン文

字で示しても良い。

4 8. 発明の名称は、要素 InventionTitle において出願言語で示さなければならず、複数の要素 InventionTitle を使用して追加言語で示しても良い（第 45 項の表を参照）。適切な言語コード（第 9 項の「国際標準化機構(ISO)言語コード 639-1:2002 年公表」の参照を参照）は、要素の属性 languageCode で示さなければならない。

4 9. 次の例は、上記の第 47 項および第 48 項による発明者名および発明の名称の表示を示す。

例：出願人名および発明者名をそれぞれ日本語とラテン文字で表記し、発明の名称を日本語、英語およびフランス語で表記する。

```
<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>
<InventionTitle languageCode="ja">efg 蛋白質をコードするマウス abcd-1 遺伝子</InventionTitle>
<InventionTitle languageCode="en">Mus musculus abcd-1 gene for efg protein</InventionTitle>
<InventionTitle languageCode="fr">Gène abcd-1 de Mus musculus pour protéine efg</InventionTitle>
```

配列データ部

5 0. 配列データ部は、1 つまたは複数の要素 SequenceData で構成されていなければならない。各要素には 1 つの配列に関する情報が含まれていなければならない。

5 1. 各要素 SequenceData には必須要素 sequenceIDNumber を有し、各配列の配列番号（第 10 項参照）が含まれていなければならない。例えば、以下ようになる。

```
<SequenceData sequenceIDNumber="1">
```

5 2. 要素 SequenceData は、次のような従属要素で構成される従属要素 INSDSeq を含まなければならない。

要素	記載	必須／任意	
		配列	意図的にスキップされた配列
INSDSeq_length	配列の長さ	必須	必須 値は不要
INSDSeq_moltype	分子の種類	必須	必須 値は不要
INSDSeq_division	特許出願と関連する配列であるという表示	「PAT」という値が付いている場合は、必須	必須 値は不要
INSDSeq_feature-table	配列の注釈を示すリスト	必須	含めては <u>ならない</u>
INSDSeq_sequence	配列	必須	必須 値に「000」を記述する。

5 3. 要素 INSDSeq_length は、要素 INSDSeq_sequence に含まれる配列のヌクレオチドまたはアミノ酸の数を開示しなければならない。例えば、次のようになる。

```
<INSDSeq_length>8</INSDSeq_length>
```

5 4. 要素 INSDSeq_moltype では、表現されている分子タイプを開示しなければならない。塩基配列（塩基類似体の配列を含む）の場合、分子タイプは DNA または RNA として示さなければならない。アミノ酸配列の場合、分子タイプは AA と表示しなければならない。（この要素は、第 55 項および第 84 項で説明している qualifier「mol_type」とは異なる）。例えば、次のようになる。

```
<INSDSeq_moltype>AA</INSDSeq_moltype>
```

5 5. 1 つまたは複数のヌクレオチドの DNA および RNA 両方のセグメントを含む塩基配列の場合、分子タイプは DNA と表示しなければならない。複合 DNA/RNA 分子は、feature key「source」、必須 qualifier「organism」の値が「synthetic construct」、必須 qualifier「mol_type」の値が「otherDNA」であることを用いて、feature table にさらに記

述しなければならない。複合 DNA/RNA 分子の DNA および RNA セグメントは、feature key「misc_feature」およびその qualifier 「note」を用いてさらに記述しなければならない。

5 6. 次の例では、上記の第 55 項に準じて、DNA セグメントおよび RNA セグメント両方を含む塩基配列の表示を示す。

```
<INSDSeq>
  <INSDSeq_length>120</INSDSeq_length>
  <INSDSeq_moltype>DNA</INSDSeq_moltype>
  <INSDSeq_division>PAT</INSDSeq_division>
  <INSDSeq_feature-table>
    <INSDFeature>
      <INSDFeature_key>source</INSDFeature_key>
      <INSDFeature_location>1..120</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>other DNA</INSDQualifier_value>
        </INSDQualifier>
      </INSDFeature_qual>
    </INSDFeature>
    <INSDFeature>
      <INSDFeature_key>misc_feature</INSDFeature_key>
      <INSDFeature_location>1..60</INSDFeature_location>
      <INSDFeature_qual>
        <INSDQualifier>
          <INSDQualifier_name>note</INSDQualifier_name>
          <INSDQualifier_value>DNA</INSDQualifier_value>
        </INSDQualifier>
      </INSDFeature_qual>
    </INSDFeature>
    <INSDFeature>
      <INSDFeature_key>misc_feature</INSDFeature_key>
      <INSDFeature_location>61..120</INSDFeature_location>
      <INSDFeature_qual>
        <INSDQualifier>
          <INSDQualifier_name>note</INSDQualifier_name>
          <INSDQualifier_value>RNA</INSDQualifier_value>
        </INSDQualifier>
      </INSDFeature_qual>
    </INSDFeature>
  </INSDSeq_feature-table>
  <INSDSeq_sequence>cgaccacgcgtccgaggaaccaaccatcagcttgaggactctgtgaaggaattggataatacccgccctaccaaagt
  gcgagcgcgactcattgctcctcgtagcgagcggc</INSDSeq_sequence>
</INSDSeq>
```

5 7. 要素 INSDSeq_sequence は、配列を開示しなければならない。附属書 1 に記載されている適切な記号（第 1 節、表 1 および第 3 節、表 3）のみを配列に含めなければならない。数字、句読点または空白文字を配列に含めてはならない。

5 8. 意図的にスキップされた配列は、配列表に記載しなければならず、次のように表示しなければならない。

- (a) 要素 SequenceData とその属性 sequenceIDNumber で、スキップされた配列の配列番号が値として提供される。

- (b) 要素 INSDSeq_length, INSDSeq_moltype, INSDSeq_division は存在するが値は提供されない。
- (c) 要素 INSDSeq_feature-table を記載してはならない。また、
- (d) 文字列「000」を値とする要素 INSDSeq_sequence

5 9. 次の例は、上記の第 58 項に従い、意図的にスキップされた配列の表示を示す。

```
<SequenceData sequenceIDNumber="3">
  <INSDSeq>
    <INSDSeq_length/>
    <INSDSeq_moltype/>
    <INSDSeq_division/>
    <INSDSeq_sequence>000</INSDSeq_sequence>
  </INSDSeq>
</SequenceData>
```

feature table

6 0. feature table には、特定の配列内の様々な領域の位置と役割に関する情報が含まれている。feature table は、意図的にスキップされた配列以外のすべての配列に必要である。feature table は 1 つまたは複数の要素 INSDFeature からなる要素 INSDSeq_feature-table に含まれる。

6 1. 各要素 INSDFeature は、1 つの feature を記述し、以下のような従属要素で構成されている。

要素	記載	必須／任意
INSDFeature_key	feature を示す文字または略語	必須
INSDFeature_location	feature に対応する配列内の領域	必須
INSDFeature_qual	feature に関する補助情報を含む qualifier	feature key が 1 つまたは複数の qualifier (例. source, otherwise, Optional)を必要とする場合は必須

feature keys

6 2. 附属書 1 には、本標準に使用されなければならない feature key の一覧と、関連する qualifier の一覧、及びこれらの qualifier が必須であるか任意であるかの指示が記載されている。附属書 1 の第 5 節には塩基配列に用いる feature key が、第 7 節にはアミノ酸配列に用いる feature key がそれぞれ掲載されている。

必須の feature keys

6 3. 意図的にスキップされた配列を除き、feature key「source」はすべての塩基配列およびすべてのアミノ酸配列に必須である。各配列には、その配列全体をその範囲とする単一の feature key「source」が必要である。1 つの配列が複数の起源に由来する場合、それらの起源は、塩基配列では feature key「misc_feature」とその qualifier「note」を、アミノ酸配列では feature key「REGION」とその qualifier「note」を用いて、feature table にさらに記述することができる。

feature location

6 4. 必須要素 INSDFeature_location は、要素 INSDSeq_sequence 内の feature に対応する部位または領域を定義する、少なくとも 1 つの存在位置記述子を含まなければならない。アミノ酸配列は、必須要素 INSDFeature_location に 1 つのみ存在位置記述子を含まなければならない。塩基配列は、1 つまたは複数の存在位置演算子と組み合わせて使用される場合、必須要素 INSDFeature_location に 1 つま

たは複数の存在位置記述子を持つことができる（第 67 項から 70 項参照）。

6 5. 存在位置記述子は、単一の残基番号、残基番号の連続した範囲を区切る領域、または特定の残基または残基の範囲を超えて広がる部位または領域とすることができる。存在位置記述子は、要素 INSDSeq_sequence の範囲を超える残基の数を含んではならない。塩基配列のみ、存在位置記述子は 2 つの隣接する残基番号間の部位とすることができる。ある配列の特徴が塩基配列の不連続な部位や領域に対応する場合には、複数の存在位置記述子を存在位置演算子と組み合わせて使用しなければならない（第 67 項から 70 項参照）。

6 6. 各タイプの存在位置記述子の構文を以下の表に示す。x と y は正の整数で示される残基番号で、要素 INSDSeq_sequence の配列の長さ以下であり、x は y 未満である。

(a) 塩基およびアミノ酸配列の存在位置記述子

存在位置記述子のタイプ	構文	記載
単一の残基番号	x	配列上の単一の残基を指す。
配列の範囲を区切る残基番号	x..y	先頭の残基と末端の残基で囲まれた、残基の連続した範囲を指す。
特定の区間の先頭の残基番号の前に存在する残基と、特定の区間の末端の残基番号の後に存在する残基	<x >x <x..y x..>y <x..>y	指定された残基または残基の範囲を含み、指定された残基を超えて広がる領域を指す。記号「<」および「>」は単一の残基、または残基の範囲の先頭および末端の残基番号とともに使用しても良く、feature が指定された残基番号を超えて広がることを示すことができる。

(b) 塩基配列のみの存在位置記述子

存在位置記述子のタイプ	構文	記載
2 つの隣接するヌクレオチド間にある部位	x^y	2 つの隣接するヌクレオチド間の部位（エンドヌクレアーゼによる切断部位など）を指す。隣接するヌクレオチドの位置番号は caret (^) で区切られている。この記述子に許される形式は、x^x+1（例えば 55^56）、または環状ヌクレオチドの場合は x^1 で、ここで「x」は分子の全長で、例えば長さ 1000 の環状分子の場合は 1000^1 となる。

(c) アミノ酸配列のみの存在位置記述子

存在位置記述子のタイプ	構文	記載
鎖内架橋により結合した残基番号	x..y	「CROSSLNK」や「DISULFID」のような鎖内架橋を示す機能と共に使用される場合、鎖内架橋によって結合されたアミノ酸を指す。

6 7. 塩基配列の要素 INSDFeature_location には、1 つまたは複数の存在位置演算子が含まれることがある。存在位置演算子は、単一だが不連続な feature に対応する 1 つの存在位置記述子または存在位置記述子の組み合わせの接頭辞であり、示された配列上の feature に対応する位置情報、またはその feature の構築方法を指定する。次に、存在位置演算子の一覧とその定義を示す。ロケーション演算子はヌクレオチドに対してのみ使用できる。

location の構文	location の記載
join(location,location, ... location)	示されている位置情報は、1 つの連続した配列になる様に結合している(端-端接合)。
order(location,location, ... location)	要素は特定の順序で表示されているが、これらの要素をまとめて表示することが適切であるかどうかに関しては示唆されていない。
complement(location)	5'位から 3'位の方向、または 5'位から 3'位の方向を模した方向に読んだ場合、存在位置記述子で指定された配列範囲に相補的な鎖上に feature が位置することを示している。

6 8. join および order 存在位置演算子では、少なくとも 2 つのコンマ区切りの存在位置記述子を指定する必要がある。2 つの隣接する残基間の部位を含む存在位置記述子 (x^y) は、join または order location 内で使用してはならない。join 存在位置演算子の使用は、存在位置記述子によって記述された残基が生物学的プロセスによって物理的に接触していることを意味する（例えば、コーディング領域の feature に寄与するエクソンなど）。

69. 存在位置演算子「complement」は、同じロケーションで「join」または「order」と組み合わせて使用できる。同じロケーションで「join」と「order」を組み合わせて使用してはならない。

70. 次の例は、上記の第 64 項から 69 項に従って、feature location の表示を示す。

(a) 塩基配列およびアミノ酸配列の location

location の例	記載
467	配列内の残基 467 を示している。
340..565	残基 340 と残基 565 に囲まれた、残基の連続した範囲を示している。
<1	先頭の残基の前の feature location を示している。
<345..500	feature の正確な下側の境界点が不明であることを示している。その location は 345 の残基の前に位置するいずれかの残基から始まって、残基 500 の位置まで（残基 500 を含む）継続していることを示している。
<1..888	先頭に配列された残基の前から始まり、残基 888 まで（残基 888 を含む）継続する feature を示している。
1..>888	先頭に配列された残基から始まり、残基 888 を超えて継続する feature を示している。
<1..>888	先頭に配列された残基の前から始まり、残基 888 を超えて継続する feature を示している。

(b) 塩基配列のみに関する location

location の例	記載
123^124	残基 123 と残基 124 間の部位を示している。
join(12..78,134..202)	領域 12~78 と 134~202 を結合して、1 つの連続した配列を形成することを示している。
complement(34..126)	ヌクレオチド 126 を補完するヌクレオチドを先頭に、ヌクレオチド 34 を補完するヌクレオチドを末端とすることを示している（feature は、残基番号が付いた本鎖を補完する鎖上に存在する）。
complement(join(2691..4571, 4918..5163))	ヌクレオチド 2691 からヌクレオチド 4571 までと、ヌクレオチド 4918 からヌクレオチド 5163 までが結合しており、feature はその結合したセグメントを補完していることを示している（feature はこの結合した鎖を補完する鎖上に存在する）。
join(complement(4918..5163), complement(2691..4571))	ヌクレオチド 4918 からヌクレオチド 5163 までの領域と、ヌクレオチド 2691 からヌクレオチド 4571 までの領域を補完しており、補完するセグメントと補完されるセグメントが結合していることを示している（feature はこの結合した鎖を補完する鎖上に存在する）。

(c) アミノ酸配列のみに関する location

location の例	記載
340..565	「CROSSLNK」や「DISULFID」のような鎖内架橋を示す feature と共に使用すると、340 位と 565 位のアミノ酸が鎖内架橋で結合していることを示している。

71. 配列表の XML インスタンスでは、存在位置記述子の文字「<」および「>」は、適切な予め定義された文字実体に置き換えなければならない（第 4 1 項参照）。例えば、次のようになる。

Feature location "<1":
<INSDFeature_location><1</INSDFeature_location>

Feature location "1..>888":
<INSDFeature_location>1..>888</INSDFeature_location>

feature qualifiers

7 2. qualifier は、feature key および feature location で伝えられる情報に加えて、feature に関する情報を提供するために使用される。qualifier によって伝えられる様々な種類の情報に対応するために、次の 3 種類の値の形式がある。

(a) フリーテキスト（第 85 項から 87 項参照）

(b) 統制語彙または計数值（例えば、数字や日付）および、

(c) 配列

7 3. 附属書 I の第 6 節では、各塩基配列の feature key に用いる qualifier とその指定された値の形式（存在する場合は）専用のリストが、第 8 節では、各アミノ酸配列の feature key に用いる qualifier 専用のリストが示されている。

7 4. 第 7 項に含まれる配列のうち、qualifier の値として提供されるものは、配列表に別途記載され、独自の配列番号が割り振られなければならない（第 10 節参照）。

必須の feature qualifiers

7 5. 必須の feature key、すなわち、塩基配列およびアミノ酸配列の「source」には、「organism」および「mol_type」という 2 つの必須 qualifier が必要である。一部の任意の feature key にも、必須 qualifier が必要である。

qualifier の要素

7 6. 要素 INSDFeature_qual は、1 つまたは複数の要素 INSDQualifier を含む。各要素 INSDQualifier は 1 つの qualifier を示しており、次の 3 つの従属要素で構成されている。

要素	記載されている情報	必須／任意
INSDQualifier_name	qualifier の名称（附属書 I、第 6 節および第 8 節参照）。	必須
INSDQualifier_value	qualifier の値がある場合は、指定された形式（附属書 I、第 6 節および第 8 節参照）で、および、第 40 項(b)に記載された文字で構成される。	指定されている場合は必須（第 87 項、附属書 I、第 6 節および第 8 節参照）
NonEnglishQualifier_value	qualifier の値がある場合には、指定された形式（附属書 I の第 6 節および第 8 節参照）で、および、第 40 項(a)に記載された文字で構成される。	指定されている場合は必須（第 87 項、附属書 I、第 6 節および第 8 節参照）

7 7. 生物に関する qualifier、すなわち塩基配列の場合は「organism」(附属書 I、第 6 節参照)、アミノ酸配列の場合は「organism」(附属書 I、第 8 節参照)は、その配列のソース、すなわち単一の生物または由来を開示しなければならない。生物名は、生物分類データベースから選択されるべきである。

7 8. 配列が自然発生的であり、由来する生物がラテン語の属名と種名を有する場合は、その名称を qualifier の値として使用しなければならない。推奨する英語の俗名は、塩基配列およびアミノ酸配列ともに qualifier「note」を用いて明示しても良いが、organism qualifier の値として使用してはならない。

7 9. 次の例では、上記の第 77 項および第 78 項に従った配列の由来を示す。

例 1：塩基配列の由来

```
<INSDSeq_feature-table>
  <INSDFeature>
    <INSDFeature_key>source</INSDFeature_key>
```

```

<INSDFeature_location>1..5164</INSDFeature_location>
<INSDFeature_qual>
  <INSDQualifier>
    <INSDQualifier_name>organism</INSDQualifier_name>
    <INSDQualifier_value>Solanum lycopersicum</INSDQualifier_value>
  </INSDQualifier>
  <INSDQualifier>
    <INSDQualifier_name>note</INSDQualifier_name>
    <INSDQualifier_value>common name: tomato</INSDQualifier_value>
  </INSDQualifier>
  <INSDQualifier>
    <INSDQualifier_name>mol_type</INSDQualifier_name>
    <INSDQualifier_value>genomic DNA</INSDQualifier_value>
  </INSDQualifier>
</INSDFeature_qual>
</INSDFeature>
</INSDSeq_feature-table>

```

例 2. アミノ酸配列の由来

```

<INSDSeq_feature-table>
  <INSDFeature>
    <INSDFeature_key>source</INSDFeature_key>
    <INSDFeature_location>1..174</INSDFeature_location>
    <INSDFeature_qual>
      <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_qual>
  </INSDFeature>
</INSDSeq_feature-table>

```

8 0. 配列が自然発生的で、その由来となる生物が既知のラテン語属であるが、種が特定されていない、または未確認である場合、その生物の qualifier 値はラテン語属の後に「sp.」を付さなければならない。例えば、以下のようになる。

```

<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Bacillus sp.</INSDQualifier_value>

```

8 1. 配列は自然発生的であるが、生物のラテン語の属名や種名が不明の場合、生物の qualifier 値は「unidentified」と表示しなければならない。既知の分類学上の情報は、塩基配列およびアミノ酸配列ともに qualifier「note」で示すべきである。例えば、次のようになる。

```

<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>unidentified</INSDQualifier_value>
<INSDQualifier_name>note</INSDQualifier_name>
<INSDQualifier_value>bacterium B8</INSDQualifier_value>

```

8 2. 配列は自然発生的であり、ウイルスのように元となる生物がラテン語の属名と種名を有しない場合は別の許容できる学名（例として、「犬伝染性肝炎 2 型」）を生物の qualifier の値として使用しなければならない。例えば、次のようになる。

```

<INSDQualifier_name>organism</INSDQualifier_name>
<INSDQualifier_value>Canine adenovirus type 2</INSDQualifier_value>

```

8 3. 配列が自然発生的で無い場合、生物の qualifier 値を「synthetic construct」と表示しなければならない。

配列の生成方法に関するさらなる情報は、塩基配列およびアミノ酸配列ともに qualifier「note」を用いて示しても良い。例えば、次のようになる。

```
<INSDSeq_feature-table>
  <INSDFeature>
    <INSDFeature_key>source</INSDFeature_key>
    <INSDFeature_location>1..40</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>
        <INSDQualifier_name>note</INSDQualifier_name>
        <INSDQualifier_value>synthetic peptide used as assay for
        antibodies</INSDQualifier_value>
      </INSDQualifier>
    </INSDFeature_qual>
  </INSDFeature>
</INSDSeq_feature-table>
```

8 4. 塩基配列の qualifier「mol_type」（附属書 1、第 6 節参照）と、アミノ酸配列の qualifier「mol_type」（附属書 1、第 8 節参照）は、配列で示される分子のタイプを開示しなければならない。これらの qualifier は、第 54 項で説明した要素 INSDSeq_moltype とは異なる。

(a) 塩基配列の場合、qualifier「mol_type」の値は、以下のいずれかでなければならない。「genomic DNA」、「genomic RNA」、「mRNA」、「tRNA」、「rRNA」、「other RNA」、「other DNA」、「transcribed RNA」、「viral cRNA」、「unassigned DNA」、または「unassigned RNA」。配列が自然発生的で無い場合は、すなわち qualifier「organism」の値が「synthetic construct」である場合、qualifier「mol_type」の値は「other RNA」または「other DNA」のいずれかでなければならない。

(b) アミノ酸配列の場合、qualifier「mol_type」の値は「protein」である。

フリーテキスト

8 5. 第 3 項(n)に記載のとおり、フリーテキストは特定の qualifier（附属書 1 の記載のとおり）値のフォーマットの一種であり、説明的なテキストフレーズの形式で提示され、望ましくは英語またはその他の指定されたフォーマット（附属書 1 記載のとおり）で提示される。

8 6. フリーテキストの使用は、配列の特性を理解するために不可欠ないくつかの短い用語に限定されていなければならない。各 qualifier について、フリーテキストの文字数は 1,000 文字を超えてはならない。

8 7. 第 3 項(o)に記載のとおり、言語に依存したフリーテキストは、国内、広域、国際的な手続きのために翻訳が必要になる可能性があるという点で、言語に依存する特定の qualifier のフリーテキストの値である。言語に依存するフリーテキスト値のフォーマットを持つ塩基配列の qualifier は、附属書 1 の第 6 節、表 5 に示されている。言語に依存したフリーテキスト形式のアミノ酸配列の qualifier は、附属書 1 の第 8 節、表 6 に記載されている。

(a) 言語に依存するフリーテキストは、英語では要素 INSDQualifier_value で、英語以外の言語では要素 NonEnglishQualifier_value で、またはその両方の要素で提示しなければならない。なお、生物名がラテン語の属名および種名である場合には、翻訳は必要ないことに留意されたい。国際的に使用されている英語以外の言葉に由来する専門用語や固有名詞は、要素 INSDQualifier_value の値の目的のために英語とみなされる（例として、「in vitro」、「in vivo」）。

(b) 要素 NonEnglishQualifier_value が配列表に存在する場合、ルート要素 nonEnglishFreeTextLanguageCode 属性に適切な言語コード（第 9 項の「国際標準化機構(ISO)言語

コード「639-1:2002 年公表」参照を参照) が示されなければならない (第 43 項参照)。単一の配列表内のすべての要素 NonEnglishQualifier_value は、nonEnglishFreeTextLanguageCode 属性で示された言語の値を持たなければならない。要素 NonEnglishQualifier_value は、言語に依存したフリーテキストの値のフォーマットを有する qualifier に対してのみ許可される。

- (c) NonEnglishQualifier_value と INSDQualifier_value が 1 つの qualifier に対して共に存在する場合、2 つの要素に含まれる情報は同等でなければならない。すなわち、NonEnglishQualifier_value に INSDQualifier_value の値の翻訳が含まれている、または、INSDQualifier_value に NonEnglishQualifier_value の値の翻訳が含まれている、または、両方の要素に originalFreeTextLanguageCode 属性で指定された言語からの qualifier 値の翻訳が含まれている、という条件のいずれかが真でなければならない (第 43 項参照)。
- (d) 言語に依存する qualifier の場合、要素 INSDQualifier はオプションの属性 ID を含んでも良い。この属性の値は、「q」の後に「q23」のような正の整数を続けた形式でなければならない、1 つの要素 INSDQualifier に固有でなければならない。つまり、属性値は、配列表ファイルで 1 度のみの使用としなければならない。

8 8. 次の例は、第 86 項の 2 で説明された言語依存のフリーテキストの提示を示す。

例 1：要素 INSDQualifier_value における言語依存のフリーテキスト

```
<INSDFeature>
  <INSDFeature_key>regulatory</INSDFeature_key>
  <INSDFeature_location>1..60</INSDFeature_location>
  <INSDFeature_qual>
    <INSDQualifier id="q1">
      <INSDQualifier_name>function</INSDQualifier_name>
      <INSDQualifier_value>binds to regulatory protein Est3</INSDQualifier_value>
    </INSDQualifier>
  </INSDFeature_qual>
</INSDFeature>
```

例 2：要素 INSDQualifier_value と要素 NonEnglishQualifier_value における言語依存のフリーテキスト

```
<INSDFeature>
  <INSDFeature_key>ACT_SITE</INSDFeature_key>
  <INSDFeature_location>51..64</INSDFeature_location>
  <INSDFeature_qual>
    <INSDQualifier id="q45">
      <INSDQualifier_name>note</INSDQualifier_name>
      <INSDQualifier_value>cleaves carbohydrate chain</INSDQualifier_value>
    <NonEnglishQualifier_value>clive la chaîne glucidique</NonEnglishQualifier_value>
  </INSDQualifier>
</INSDFeature_qual>
</INSDFeature>
```

例 3：要素 NonEnglishQualifier_value における言語依存のフリーテキスト。

```
<INSDFeature>
  <INSDFeature_key>ACT_SITE</INSDFeature_key>
  <INSDFeature_location>51..64</INSDFeature_location>
  <INSDFeature_qual>
    <INSDQualifier id="q1034">
      <INSDQualifier_name>note</INSDQualifier_name>
      <NonEnglishQualifier_value>clive la chaîne glucidique </NonEnglishQualifier_value>
    </INSDQualifier>
  </INSDFeature_qual>
</INSDFeature>
```

コード配列

89. feature key「CDS」はコード配列、すなわち蛋白質のアミノ酸の配列に対応する塩基配列と終止コドン
を識別するために使用しても良い。必須要素 INSDFeature_location における feature「CDS」location に
は、終止コドンが含まれていなければならない。
90. qualifier「transl_table」および「translation」は、feature key「CDS」と共に用いても良い（附属書 I 参
照）。qualifier「transl_table」を使用しない場合は、標準コード表（附属書 I、第 9 節、表 5 参照）の使
用が想定される。
91. qualifier「transl_except」は、ピロリジン、またはセレノシステインのいずれかをコード化するコドン
を識別するために、feature key「CDS」および qualifier「translation」と共に使用しなければならない。
92. 第 7 項に包含される qualifier「translation」に開示され、コード配列によってコードされたアミノ酸
配列は、配列表に記載され、個別の配列番号が割り当てられなければならない。アミノ酸配列
に割り当てられた配列番号は、feature key「CDS」を有する qualifier「protein_id」の値として記載
されなければならない。アミノ酸配列に用いる feature key「source」の qualifier「organism」は、
そのコード配列のものと同じでなければならない。例えば、次のようになる。

```
<INSDFeature>
```

```
  <INSDFeature_key>CDS</INSDFeature_key>
```

```
  <INSDFeature_location>1..507</INSDFeature_location>
```

```
  <INSDFeature_qual>
```

```
    <INSDQualifier>
```

```
      <INSDQualifier_name>transl_table</INSDQualifier_name>
```

```
      <INSDQualifier_value>11</INSDQualifier_value>
```

```
    </INSDQualifier>
```

```
    <INSDQualifier>
```

```
      <INSDQualifier_name>translation</INSDQualifier_name>
```

```
      <INSDQualifier_value>MLVHLERTTIMFDFSSLINLPLIWGLLIAlAVLLYLMDGFDLGI  
GILLPFAPSDKCRDHMISSIAPFWDGNETWLVLGGGGLFAAFPLAYSILMPAFYIPIIIM  
LLGLIVRGVSFEFRFKAEGKYRRLWDYAFHFGSLGAAFCQGMILGAFIHGVEVNGRN  
FSGGQLM
```

```
    </INSDQualifier_value>
```

```
  </INSDQualifier>
```

```
  <INSDQualifier>
```

```
    <INSDQualifier_name>protein_id</INSDQualifier_name>
```

```
    <INSDQualifier_value>89</INSDQualifier_value>
```

```
  </INSDQualifier>
```

```
</INSDFeature_qual>
```

```
</INSDFeature>
```

変異体

93. 一次配列とその変異体配列、それぞれがそれらの残基の数え上げにより開示され、第 7 項に含まれるも
のは、それぞれ配列表に含め、個別の配列番号を割り当てなければならない。
94. 1 つまたは複数の位置に数え上げられた代替残基を有する単一配列として開示された任意の変異体配列
は、配列表に記載しなければならず、連続した代替残基が最も限定的な曖昧さ記号で表された単一配列
で表示すべきである（第 15 項および第 27 項を参照のこと）。
95. 一次配列における欠失、挿入または置換によってのみ開示される任意の変異体配列は、配列表に記載す
べきである。配列表に含まれる場合、そのような変異体配列は、
- 単一の location または複数の異なる location に変異があり、それらの変異の発生が独立している場
合、一次配列の注釈を付して表しても良い。
 - 複数の異なる location に変異があり、それらの変異の発生が相互に依存している場合には、別の配
列として表現し、個別の配列番号を割り当てるべきである。

(c) 1,000 を超える残基を含む挿入または置換された配列を含む場合は、別の配列として表現し、個別の配列番号を割り振るべきである（第 86 項を参照のこと）。

9 6. 次の表は、核酸およびアミノ酸配列の変異体の feature key と qualifier の適切な使用方法を示している。

配列の種類	feature key	qualifier	利用
核酸	variation	replace または note	自然発生の変異および多型、例. 対立遺伝子、制限酵素セグメント長多型(RFLP)
核酸	misc_difference	replace または note	人工的に導入された多様性、例. 遺伝子操作または化学合成
アミノ酸	VAR_SEQ	note	選択的スプライシング、プロモーターの選択的利用、新規転写の開始およびリボソームフレームシフトにより生じる変異体
アミノ酸	VARIANT	note	「VAR_SEQ」が使用できない任意の変異体の種類

9 7. 上の表に示されているように、特定の変異体に対する配列の注釈には、feature key、qualifier および feature location を含まなければならない。qualifier「replace」の値は、第 1 節、表 1 に記載された記号のみを使用した単一の代替ヌクレオチドまたは塩基配列のみでなければならない、または空でなければならない。qualifier「note」の値として、代替残基のリストを記載しても良い。特に、代替アミノ酸のリストは、qualifier「note」の値として記載しなければならない。ここで、配列に「X」が用いられており、

「X」は「A」、「R」、「N」、「D」、「C」、「Q」、「E」、「G」、「H」、「I」、「L」、「K」、「M」、「F」、「P」、「O」、「S」、「U」、「T」、「W」、「Y」、または「V」のいずれか 1 つ以外の値を示す(第 27 項参照)。欠失している場合は、qualifier「replace」の値を空白にするか、qualifier「note」にその残基が欠失している可能性があることを示す表示をしなければならない。挿入または置換された残基は、qualifier「replace」または qualifier「note」で提示されなければならない。qualifier「replace」および「note」の値の形式は、第 86 項で規定されているように、フリーテキストであり、1,000 文字を超えてはならない。qualifier の値に挿入または置換として記載されている第 7 項に記載されている配列については、第 100 項を参照のこと。

9 8. 附属書 I に記載されている記号（第 1 節から 4 節、表 1 から 4 をそれぞれ参照）は、必要に応じて変異体の残基を表示するために使用されるべきである。qualifier「note」に関しては、変異体の残基が附属書 I の表 2 または表 4 に記載されていない修飾残基である場合、qualifier 値として修飾残基の省略されていない完全な名称を記載しなければならない。修飾残基は、第 17 項または第 30 項に記載があるように、feature table に詳細に記載しなければならない。

9 9. 次の例では、上記の第 95 項から 98 項に準拠型の変異体の表示を示す。

例 1：数え上げられた代替ヌクレオチドに用いる feature key「misc_difference」。
配列の 53 位の「n」は、5 つの代替ヌクレオチドのうちの 1 つである。

```
<INSDFeature>
  <INSDFeature_key>misc_difference</INSDFeature_key>
  <INSDFeature_location>53</INSDFeature_location>
  <INSDFeature_qual>
    <INSDQualifier>
      <INSDQualifier_name>note</INSDQualifier_name>
      <INSDQualifier_value>w, cmnm5s2u, mam5u, mcm5s2u, or p</INSDQualifier_value>
    </INSDQualifier>
  </INSDFeature_qual>
</INSDFeature>
<INSDFeature>
  <INSDFeature_key>modified_base</INSDFeature_key>
  <INSDFeature_location>53</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>cmnm5s2u, mam5u, mcm5s2u, or p</INSDQualifier_value>
  </INSDQualifier>
</INSDFeature_qual>
</INSDFeature>

```

例 2 : 塩基配列における欠失に用いる feature key「misc_difference」。
配列の 413 位のヌクレオチドが欠失している。

```

<INSDFeature>
  <INSDFeature_key>misc_difference</INSDFeature_key>
  <INSDFeature_location>413</INSDFeature_location>
  <INSDFeature_qual>
    <INSDQualifier>
      <INSDQualifier_name>replace</INSDQualifier_name>
      <INSDQualifier_value></INSDQualifier_value>
    </INSDQualifier>
  </INSDFeature_qual>
</INSDFeature>

```

例 3 : 塩基配列における挿入に用いる feature key「misc_difference」。
一次配列の 100 位と 101 位の間に、配列「atgccaatat」が挿入されている。

```

<INSDFeature>
  <INSDFeature_key>misc_difference</INSDFeature_key>
  <INSDFeature_location>100^101</INSDFeature_location>
  <INSDFeature_qual>
    <INSDQualifier>
      <INSDQualifier_name>replace</INSDQualifier_name>
      <INSDQualifier_value>atgccaatat</INSDQualifier_value>
    </INSDQualifier>
  </INSDFeature_qual>
</INSDFeature>

```

例 4 : 塩基配列における置換に用いる feature key「variation」。
シトシンが配列の 413 位のヌクレオチドに置換している。

```

<INSDFeature>
  <INSDFeature_key>variation</INSDFeature_key>
  <INSDFeature_location>413</INSDFeature_location>
  <INSDFeature_qual>
    <INSDQualifier>
      <INSDQualifier_name>replace</INSDQualifier_name>
      <INSDQualifier_value>c</INSDQualifier_value>
    </INSDQualifier>
  </INSDFeature_qual>
</INSDFeature>

```

例 5 : アミノ酸配列において置換に用いる feature key「VARIANT」。
配列の 100 位のロケーションにあるアミノ酸は、I, A, F, Y, alle, Melle または Nie で置換することができる。

```

<INSDFeature>
  <INSDFeature_key>VARIANT</INSDFeature_key>
  <INSDFeature_location>100</INSDFeature_location>
  <INSDFeature_qual>
    <INSDQualifier>

```

```

        <INSDQualifier_name>note</INSDQualifier_name>
        <INSDQualifier_value>I, A, F, Y, alle, Melle, or Nle
    </INSDQualifier_value>
    </INSDQualifier>
</INSDFeature_qual>
</INSDFeature>
<INSDFeature>
    <INSDFeature_key>MOD_RES</INSDFeature_key>
    <INSDFeature_location>100</INSDFeature_location>
    <INSDFeature_qual>
        <INSDQualifier>
            <INSDQualifier_name>note</INSDQualifier_name>
            <INSDQualifier_value>alle, Melle, or Nle</INSDQualifier_value>
        </INSDQualifier>
    </INSDFeature_qual>
</INSDFeature>

```

例 6 : アミノ酸配列において置換に用いる feature key「VARIANT」。
配列の 100 位に指定されたアミノ酸は、Lys, Arg または His を除くすべてのアミノ酸で置換することができる。

```

<INSDFeature>
    <INSDFeature_key>VARIANT</INSDFeature_key>
    <INSDFeature_location>100</INSDFeature_location>
    <INSDFeature_qual>
        <INSDQualifier>
            <INSDQualifier_name>note</INSDQualifier_name>
            <INSDQualifier_value>not K, R, or H</INSDQualifier_value>
        </INSDQualifier>
    </INSDFeature_qual>
</INSDFeature>

```

1 0 0. 第 7 項に含まれる配列で、一次配列の注釈の qualifier 値の挿入または置換として記載されるものも配列表に記載され、個別の配列番号が割り当てられなければならない。

[次頁から附属書 I が続く]

目 次

- 第 1 節：ヌクレオチドのリスト
- 第 2 節：修飾塩基のリスト
- 第 3 節：アミノ酸のリスト
- 第 4 節：修飾アミノ酸のリスト
- 第 5 節：塩基配列の feature keys
- 第 6 節：塩基配列の qualifiers
- 第 7 節：アミノ酸配列の feature keys
- 第 8 節：アミノ酸配列の qualifiers
- 第 9 節：遺伝コード表

第 1 節：ヌクレオチドのリスト

配列表で使用されるヌクレオチド塩基の記号を表 1 に示す。記号「t」は、詳細な説明なしに使用される場合、DNA ではチミン、RNA ではウラシルと解釈される。曖昧さ記号（2 つ以上の塩基を表す）が適切である場合は、最も限定的な記号を用いるべきである。例えば、ある位置の塩基が「a または g」である可能性がある場合、「n」ではなく「r」を使用すべきである。記号「n」は、それ以上の記述がない状態で使用された場合、「a または c または g または t/u」と解釈される。

表 1：ヌクレオチド記号のリスト

Symbol	Definition
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”

第 2 節：修飾塩基のリスト

表 2 で示されている略語は、qualifier「mod_base」に許される唯一の値である。特定の修飾されたヌクレオチドが以下の表に存在しない場合は、その値として「OTHER」という略語を使用しなければならない。省略形が「OTHER」の場合は、修飾塩基の省略されていない完全な名称を qualifier 「note」で指定しなければならない。表 2 に示されている略語は、配列自体に使用してはならない。

表 2：修飾塩基のリスト

Abbreviation	Definition
ac4c	4-acetylcytidine
chm5u	5-(carboxyhydroxymethyl)uridine
cm	2'-O-methylcytidine
cmnm5s2u	5-carboxymethylaminomethyl-2-thiouridine
cmnm5u	5-carboxymethylaminomethyluridine
dhu	dihydrouridine
fm	2'-O-methylpseudouridine
gal q	beta-D-galactosylqueuosine
gm	2'-O-methylguanosine
i	inosine
i6a	N6-isopentenyladenosine
m1a	1-methyladenosine
m1f	1-methylpseudouridine
m1g	1-methylguanosine
m1i	1-methylinosine
m22g	2,2-dimethylguanosine
m2a	2-methyladenosine
m2g	2-methylguanosine
m3c	3-methylcytidine
m4c	N4-methylcytosine
m5c	5-methylcytidine
m6a	N6-methyladenosine
m7g	7-methylguanosine
mam5u	5-methylaminomethyluridine
mam5s2u	5-methylaminomethyl-2-thiouridine
man q	beta-D-mannosylqueuosine
mcm5s2u	5-methoxycarbonylmethyl-2-thiouridine
mcm5u	5-methoxycarbonylmethyluridine
mo5u	5-methoxyuridine
ms2i6a	2-methylthio-N6-isopentenyladenosine
ms2t6a	N-((9-beta-D-ribofuranosyl-2-methylthiopurine-6-yl)carbamoyl)threonine
mt6a	N-((9-beta-D-ribofuranosylpurine-6-yl)N-methyl-carbamoyl)threonine
mv	uridine-5-oxoacetic acid-methylester
o5u	uridine-5-oxoacetic acid (v)
osyw	wybutoxosine
p	pseudouridine
q	queuosine
s2c	2-thiocytidine
s2t	5-methyl-2-thiouridine
s2u	2-thiouridine

Abbreviation	Definition
s4u	4-thiouridine
m5u	5-methyluridine
t6a	N-((9-beta-D-ribofuranosylpurine-6-yl)carbamoyl)threonine
tm	2'-O-methyl-5-methyluridine
um	2'-O-methyluridine
yw	wybutosine
x	3-(3-amino-3-carboxypropyl)uridine, (acp3)u
OTHER	(requires note qualifier)

第 3 節：アミノ酸のリスト

配列に使用するアミノ酸記号を表 3 に示す。曖昧さ記号（2 つ以上のアミノ酸を表す）が適切な場合は、最も限定的な記号を使用すべきである。例えば、ある位置のアミノ酸がアスパラギン酸またはアスパラギンである可能性がある場合には、「X」ではなく「B」という記号を使用すべきである。記号「X」は、それがさらなる説明なしに使用される場合、「A」、「R」、「N」、「D」、「C」、「Q」、「E」、「G」、「H」、「I」、「L」、「K」、「M」、「F」、「P」、「O」、「S」、「U」、「T」、「W」、「Y」、または「V」のいずれか 1 つと解釈される。

表 3：アミノ酸記号のリスト

Symbol	Definition
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"

第 4 節：修飾アミノ酸のリスト

表 4 は、feature key「MOD__RES」または「SITE」の必須 qualifier「note」において、修飾アミノ酸に許される唯一の略語の一覧である。qualifier「note」の値は、必要に応じてこの表の略語か、修飾アミノ酸の省略されていない完全な名称のいずれかでなければならない。この表に記載されている略語(または完全な名称)は、配列自体に使用してはならない。

表 4：修飾アミノ酸のリスト

Abbreviation	Modified Amino acid
Aad	2-Aminoadipic acid
bAad	3-Aminoadipic acid
bAla	beta-Alanine, beta-Aminopropionic acid
Abu	2-Aminobutyric acid
4Abu	4-Aminobutyric acid, piperidinic acid
Acp	6-Aminocaproic acid
Ahe	2-Aminoheptanoic acid
Aib	2-Aminoisobutyric acid
bAib	3-Aminoisobutyric acid
Apm	2-Aminopimelic acid
Dbu	2,4-Diaminobutyric acid
Des	Desmosine
Dpm	2,2'-Diaminopimelic acid
Dpr	2,3-Diaminopropionic acid
EtGly	N-Ethylglycine
EtAsn	N-Ethylasparagine
Hyl	Hydroxylysine
aHyl	allo-Hydroxylysine
3Hyp	3-Hydroxyproline
4Hyp	4-Hydroxyproline
Ide	Isodesmosine
alle	allo-Isoleucine
MeGly	N-Methylglycine, sarcosine
Melle	N-Methylisoleucine
MeLys	6-N-Methyllysine
MeVal	N-Methylvaline
Nva	Norvaline
Nle	Norleucine
Orn	Ornithine

	コメント	feature「centromere」は、染色体が保持され、動原体が形成される領域に対応する DNA の区間を示すものである。
5.4.	feature key 定義 任意の qualifier	D-loop 置換ループ：ミトコンドリア DNA 内の二本鎖の一方に相同な配列の一本鎖が対合をなす領域；RecA 蛋白質の触媒反応において、二本鎖 DNA の一方が外来の一本鎖によって置換する部位の記述にも用いられる。 allele gene gene_synonym map note
	分子の分類	DNA
5.5.	feature key 定義 任意の qualifier	D_segment 免疫グロブリン重鎖および t-細胞受容体のβ鎖の多様性セグメント。 allele gene gene_synonym map note product pseudo pseudogene standard_name
	生物の分類	真核生物
5.6.	feature key 定義 任意の qualifier	exon スプライシング後の mRNA、rRNA、tRNA の一部をコード化するゲノムの領域；5'非翻訳領域、すべての CDS、3'非翻訳領域を含む場合がある。 allele EC_number function gene gene_synonym map note number product pseudo pseudogene standard_name trans_splicing
5.7.	feature key 定義 任意の qualifier	gene 遺伝子として特定され、名称が付された生物学的な関心領域。 allele function gene gene_synonym map note operon pseudo pseudogene phenotype standard_name trans_splicing
	コメント	feature「gene」は、遺伝的形質や表現型に対応する DNA の区間を記述する。feature は、定義上、末端の位置に厳密に拘束されるものではない；遺伝子が位置する領域を表すことを意味する。
5.8.	feature key 定義 任意の qualifier	iDNA DNA の介在配列；数種類の組み換えにより排除された DNA。 allele

		function gene gene_synonym map note number standard_name
	分子の分類 コメント	DNA 例、免疫グロブリン遺伝子の体細胞処理において。
5.9.	feature key 定義 任意の qualifier	intron 転写されるが、その配列（エクソン）の両端のどちらか片側をスプライシングすることにより転写産物から取り除かれる DNA のセグメント。 allele function gene gene_synonym map note number pseudo pseudogene standard_name trans_splicing
5.10.	feature key 定義 任意の qualifier	J_segment 免疫グロブリンの軽鎖と重鎖、および t-細胞受容体の α 、 β および γ 鎖の結合部。 allele gene gene_synonym map note product pseudo pseudogene standard_name
	生物の分類	真核生物
5.11.	feature key 定義 任意の qualifier	mat_peptide 成熟ペプチドまたは蛋白質のコード配列；翻訳後修飾を受けて成熟した、または最終的なペプチドまたは蛋白質産物のコード配列；その存在位置は（対応する CDS とは異なり）終止コドンを含まない。 allele EC_number function gene gene_synonym map note product pseudo pseudogene standard_name
5.12.	feature key 定義 必須の qualifier 任意の qualifier	misc_binding 他の結合 key（「primer_bind」または「protein_bind」）では表現できない別の部位を共有結合または非共有結合する核酸中の部位。 bound_moiety allele function gene gene_synonym map note

	コメント	リボソームの結合部位を記載するためには、feature key「regulatory」およびその qualifier 「regulatory_class」と共に、qualifier 値「ribosome_binding_site」を用いなければならない。
5.13.	feature key 定義 任意の qualifier	misc_difference 提示された配列とこの位置で示された配列が異なり、他の Difference キー（「variation」「modified_base」）では表現できない。 allele flone compare gene gene_synonym map note phenotype replace standard_name
	コメント	Feature key「misc_difference」は人為的に導入された変異性（例. 遺伝子操作または化学合成）に関する記載を行うために用いなければならない。欠失、挿入または置換に関する注釈を付すために qualifier「replace」を用いる。feature key「variation」は自然発生的な遺伝的変異を記述するために用いなければならない。
5.14.	feature key 定義 任意の qualifier	misc_feature 他の feature key では表現することができない生物学的な関心領域；新規または希少な feature。 allele function gene gene_synonym map note number phenotype product pseudo pseudogene standard_name
	コメント	この key は、単にある領域にコメントするため、または、他の feature location で使用するために用いるべきではない。
5.15.	feature key 定義 任意の qualifier	misc_recomb 他の組換え key や source key の qualifier「proviral」では説明できない二重鎖 DNA の切断と再結合が行われる、一般的、部位特異的、または複製的な組み換えイベントの部位（プロウイルスの）。 allele gene gene_synonym map note recombination_class standard_name
	分子の分類	DNA
5.16.	feature key 定義 任意の qualifier	misc_RNA 他の RNA キーでは定義できない転写産物や RNA 産物（prim_transcript、precursor_RNA、mRNA、5'非翻訳領域、3'非翻訳領域、exon、CDS、sig_peptide、transit_peptide、mat_peptide、intron、polyA_site、ncRNA、rRNA、tRNA）。 allele function gene gene_synonym map note operon

		product pseudo pseudogene standard_name trans_splicing
5.17.	feature key 定義 任意の qualifier	misc_structure および他の構造(Structure)キー (stem_loop、D-loop) では記述できない二次、三次のヌクレオチド構造や立体配座。 allele function gene gene_synonym map note standard_name
5.18.	feature key 定義 必須の qualifier 任意の qualifier	mobile_element 可動遺伝子因子を含む、ゲノムの領域。 mobile_element_type allele function gene gene_synonym map note rpt_family rpt_type standard_name
5.19.	feature key 定義 必須の qualifier 任意の qualifier コメント	modified_base 示されたヌクレオチドは修飾塩基で、示された分子で置換される (qualifier「mod_base」の値として記載)。 mod_base allele frequency gene gene_synonym map note 必須 qualifier「mod_base」の値は、本附属書 1、第 2 節の修飾塩基の略語として示されている統制語彙のみを用いる。
5.20.	feature key 定義 任意の qualifier	mRNA メッセンジャーRNA ; 5'非翻訳領域、コード配列(CDS, エクソン)および、3'非翻訳領域を含む。 allele circular_RNA function gene gene_synonym map note operon product pseudo pseudogene standard_name trans_splicing
5.21.	feature key 定義 必須の qualifier 任意の qualifier	ncRNA rRNA および tRNA 以外の非蛋白質コーディング遺伝子で、その機能分子が RNA 転写物であるもの。 ncRNA_class allele

		function gnene gene_synonym map note operaon product pseudo pseudogene standard_name trans_splicing
	コメント	feature 「ncRNA」は rRNA および tRNA の注釈に用いてはならない。feature key「rRNA」および feature key「tRNA」をそれぞれ用いなければならない。
5.22.	feature key 定義 任意の qualifier	N_region 再配列された免疫グロブリンセグメントの間に挿入された余分なヌクレオチド。 allele gene gene_synonym map note product pseudo pseudogene standard_name
	生物の分類	真核生物
5.23.	feature key 定義 必須の qualifier 任意の qualifier	operon 同じ制御配列/プロモーターの制御下にあり、同じ生物学的経路にある遺伝子群を含む多シストロン性転写物を含む領域。 operon allele function map note phenotype pseudo pseudogene standard_name
5.24.	feature key 定義 任意の qualifier	oriT 転写の起点；接合または可動化の過程で転写が開始される DNA 分子の領域。 allele bound_moiety direction gene gene_synonym map note rpt_family rpt_type rpt_unit_range rpt_unit_seq standard_name
	分子の分類 コメント	DNA 「rep_origin」は転写の起点を記載するために用いなければならない；その qualifier 「direction」の値として用いるのが許容される値には left、right および both があるが、feature 「oriT」と共に用いる場合は left または right のみが有効である。；転写の起点は染色体の中に存在する場合がある。；プラスミドには複数の転写の起点が存在する場合がある。
5.25.	feature key 定義 任意の qualifier	polyA_site 転写後ポリアデニル化によりアデニン残基が付加される、RNA 転写物上の部位。 allele gene

		gene_synonym map note
	生物の分類	真核生物および真核生物ウイルス
5.26.	feature key 定義 任意の qualifier	precursor_RNA 未成熟な RNA 産物ではない RNA 種；ncRNA、rRNA、tRNA、5'非翻訳領域、コード配列(CDS, エキソン)、介在配列（イントロン） および、3'非翻訳領域を含んでいても良い。 allele function gene gene_synonym map note operon product standard_name trans_splicing
	コメント	転写後のプロセシングの結果であると思われる RNA に使用される；問題の RNA が、プロセシングされていないのが知られている場合は、key「prim_transcript」を用いる。
5.27.	feature key 定義 任意の qualifier	prim_transcript (最初の、未転写の)一次転写物；ncRNA、rRNA、tRNA、5'非翻訳領域、コード配列(CDS, エキソン)、介在配列（イントロン） および、3'非翻訳領域を含んでいても良い。 allele function gene gene_synonym <u>map</u> <u>note</u> operon standard_name
5.28.	feature key 定義 任意の qualifier コメント	primer_bind 複製、転写または逆転写を開始するための非共有結合型プライマー結合部位；合成用（例. PCR プライマー要素）の部位を含む。 allele gene gene_synonym map note standard_name プライマー分子が結合する、配列上のある部位の注釈情報の記載に用いる feature key - プライマー分子そのものの配列を示すためには用いない；ポリメラーゼ連鎖(PCR)反応ではプライマー対を使用することが多いため、1つのkey「primer_bind」で2つの存在位置を指定してorder(存在位置、存在位置)演算子を用いたり、key「primer_bind」を2つ用いても良い。
5.29.	feature key 定義 任意の qualifier	propeptide プロペプチドコード配列；プロタンパク質のドメインのコード配列で、切断されて成熟蛋白質産物を形成するもの。 allele function gene gene_synonym map note product pseudo pseudogene standard_name
5.30.	feature key 定義 必須の qualifier 任意の qualifier	protein_bind 核酸上の、非共有の蛋白質結合部位。 bound_moiety allele

		function gene gene_synonym map note operon standard_name
	コメント	リボソームの結合部位を記述するには、feature key「regulatory」およびその qualifier 「regulatory_class」(qualifier 値は「ribosome_binding_site」)を用いなければならない。
5.31.	feature key 定義 必須の qualifier 任意の qualifier	regulatory 転写、翻訳、複製またはクロマチン構造の制御に機能する配列の任意の領域。 regulatory_class allele bound_moiety function gene gene_synonym map note operon phenotype pseudo pseudogene standard_name
5.32.	feature key 定義 任意の qualifier	repeat_region 繰り返し単位を含むゲノム領域。 allele function gene gene_synonym map note rpt_family rpt_type rpt_unit_range rpt_unit_seq satellite standard_name
5.33.	feature key 定義 任意の qualifier	rep_origin 複製の起点；核酸の 2 つの同じ複製を作成するための複製の起点。 allele direction function gene gene_synonym map note standard_name
	コメント	qualifier「direction」は 3 つの有効値を用いる。left、right または both。
5.34.	feature key 定義 任意の qualifier	rRNA 成熟したリボソーム RNA；アミノ酸を蛋白質に合成するリボ核蛋白質粒子（リボソーム）を構成する RNA 成分。 allele function gene gene_synonym map note operon product

		pseudo pseudogene standard_name
	コメント	qualifier 「product」を用いて rRNA の大きさに関する注釈を付すべきである。
5.35.	feature key 定義	S_region 免疫グロブリンの重鎖のスイッチ領域；同じ B 細胞から異なる免疫グロブリンクラスを発現させる重鎖 DNA の再編成に参与する。
	任意の qualifier	allele gene gene_synonym map note product pseudo pseudogene standard_name
	生物の分類	真核生物
5.36.	feature key 定義	sig_peptide シグナルペプチドのコード配列；分泌蛋白質の N 末端ドメインのコード配列；このドメインは、新生ポリペプチドを膜のリーダー配列に付着させることに参与する。
	任意の qualifier	allele function gene gene_synonym map note product pseudo pseudogene standard_name
5.37.	feature key 定義	source この key は必須；すべての配列は、配列全体を網羅する単一のソース(source) key を有する。
	必須の qualifier	organism mol_type
	任意の qualifier	cell_line cell_type chromosome clone clone_lib collected_by collection_date cultivar dev_stage ecotype environmental_sample germline haplogroup haplotype host identified_by isolate isolation_source lab_host lat_lon macronuclear map mating_type note organelle

		PCR_primers plasmid pop_variant proviral rearranged segment serotype serovar sex strain sub_clone sub_species sub_strain tissue_lib tissue_type variety すべて
5.38.	feature key 定義 任意の qualifier	stem_loop ヘアピン構造；RNA または DNA の一本鎖の中で、隣接する（逆）相補的な配列同士が塩基対合して形成された二重螺旋状の領域。 allele function gene gene_synonym map note operon standard_name
5.39.	feature key 定義 任意の qualifier 分子の分類 コメント	STS 配列タグ部；ゲノム状のマッピングランドマークを特徴づける短い、単一コピーの DNA 配列で、PCR によって検出される；一連の配列タグ部(STS)の順序を決定することでゲノムの一領域をマッピングできる。 allele gene gene_synonym map note standard_name DNA key「primer_bind」または複数のプライマーに一つまたは複数のプライマーを含めるための配列タグ部(STS)の存在位置。
5.40.	feature key 定義 任意の qualifier コメント	telomere テロメアとして特定され、実験的に特徴づけられた生物学的な関心領域。 note rpt_type rpt_unit_range rpt_unit_seq standard_name feature「telomere」は、真核生物の直線状の染色体の末端にある特定の構造に対応する DNA の区間のことで、末端の完全性と維持に必要とされる；この領域は、他の染色体と比べて特異であり、染色体の物理的な末端を示している。
5.41.	feature key 定義 任意の qualifier	tmRNA 転移伝令 RNA (tmRNA)；tmRNA は、まず tRNA として働き、次にペプチドタグをコードする mRNA として働く；リボソームが tmRNA の mRNA 領域を翻訳し、コードされたペプチドタグを未成熟蛋白質の C 末端に結合することで、このタグが蛋白質の破壊や蛋白質分解の標的となる。 allele function gene gene_synonym

		map note product pseudo pseudogene standard_name tag_peptide
5.42.	feature key 定義 任意の qualifier	transit_peptide ペプチドのコード配列：核にコードされたオルガネラ蛋白質の N 末端のドメインのコード配列；このドメインは、オルガネラへの蛋白質の翻訳後移入に参与する allele function gene gene_synonym map note product pseudo pseudogene standard_name
5.43.	feature key 定義 任意の qualifier	tRNA 成熟した tRNA：核酸配列からアミノ酸配列への翻訳を媒介する低分子 RNA（長さ 75～85 塩基）。 allele circular_RNA anticodon function gene gene_synonym map note operon product pseudo pseudogene standard_name trans_splicing
5.44.	feature key 定義 任意の qualifier コメント	unsure 配列された塩基のうち、通常 10 個以下の長さの小さな領域で、明確に特定できないもの。このような領域には、呼び出された塩基（a、t、g、c）や、呼び出された塩基と呼び出されていない塩基（「n」）が混在していることがある。 allele compare gene gene_synonym map note replace
5.45.	feature key 定義 任意の qualifier	V_region 免疫グロブリンの軽鎖、重鎖、T 細胞受容体の α 、 β 、 γ 鎖の可変領域で、可変アミノ末端部分をコードし、V_segments、D_segments、N_region、J_segments で構成されることが可能。 allele gene gene_synonym map note product

		pseudo pseudogene standard_name 真核生物
5.46.	生物の分類 feature key 定義 任意の qualifier	V_segment 免疫グロブリンの軽鎖、重鎖、T細胞受容体のα鎖、β鎖、γ鎖の可変セグメントで、可変領域（V_region）の大部分とリーダーペプチドの末端の数個のアミノ酸をコードする。 allele gene gene_synonym map note product pseudo pseudogene standard_name 真核生物
5.47.	生物の分類 feature key 定義 任意の qualifier コメント	variation 近親株に、同じ遺伝子からの不可逆の突然変異体（制限酵素セグメント長多型(RFLP)、多型など）があり、この存在位置（および他の存在位置）で提示された配列と異なる場合。 allele compare frequency gene gene_synonym map note phenotype product replace standard_name 対立遺伝子（アレル）、制限酵素セグメント長多型(RFLP)、および他の自然発生の突然変異体および多型について記載するために用いる。欠失、挿入、置換を示す注釈を付すためには qualifier「replace」を用いる；遺伝子操作（例、部位具体的変異誘発(SDM)）の結果生じる変異性は、feature key「misc_difference」を用いて記載しなければならない。
5.48.	feature key 定義 任意の qualifier コメント	3'UTR 1) 蛋白質に翻訳されない成熟した転写物の 3'末端（終止コドンの後）の領域。 2) RNA ウイルスの 3'末端の領域（最後の終止コドンの後）で、蛋白質に翻訳されない領域。 allele function gene gene_synonym map note standard_name trans_splicing アポストロフィー文字（'）は、XML 形式の文書では特別の意味があり、要素の値として記載する際は「'」で置き換えなければならない。したがって、「3'非翻訳領域」は XML ファイルでは「3'UTR」、つまり<INSDFeature_key>3'UTR</INSDFeature_key>と記載しなければならない。
5.49.	feature key 定義 任意の qualifier	5'UTR 1) タンパク質に翻訳されない成熟転写産物の 5'末端（開始コドンの前）の領域。 2) タンパク質に翻訳されない RNA ウイルスの 5'末端の領域（第 1 開始コドンの前）。 allele function gene gene_synonym map

コメント	<p>note standard_name trans_splicing</p> <p>アポストロフィー文字('')は、XML 形式の文書では特別の意味があり、要素の値として記載する際は「&apos;」と置き換えなければならない。したがって、「5'非翻訳領域」は XML ファイルでは「5&apos;UTR」、つまり<INSDFeature_key>5&apos;UTR</INSDFeature_key>と記載しなければならない。</p>
------	--

第 6 節 塩基配列の qualifiers

本節では、塩基配列の feature に使用される qualifier のリストを掲載している。qualifier はアルファベット順に並んでいる。

qualifier 値のフォーマットが「none」の場合、要素「INSDQualifier_value」は使用してはならず、要素 NonEnglishQualifier_value も使用してはならない。

値のフォーマットが言語に依存するフリーテキストの場合、以下のいずれかを使用しなければならない。

- 1) 要素 INSDQualifier_value、または
- 2) 要素 NonEnglishQualifier_value、または
- 3) 要素 INSDQualifier_value と要素 NonEnglishQualifier_value の両方。

値のフォーマットが「none」以外で、言語依存のフリーテキストとして識別されない場合、要素 INSDQualifier_value を使用しなければならない、要素 NonEnglishQualifier_value は使用してはならない。

注意：言語に依存する「フリーテキスト」値形式を有する qualifier に対して提供される qualifier 値は、国内または広域的な手続きのために翻訳が必要になる場合がある。次の表に示す qualifier は、言語に依存したフリーテキストの値を持つとみなされる。

表 5：言語に依存するフリーテキスト値を持つ塩基配列の qualifier のリスト

6.3	bound_moiety
6.5	cell_type
6.8	clone
6.9	clone_lib
6.11	collected_by
6.14	cultivar
6.15	dev_stage
6.18	ecotype
6.22	function
6.24	gene_synonym
6.26	haplogroup
6.28	host
6.29	identified_by
6.30	isolate
6.31	isolation_source
6.32	lab_host
6.36	mating_type
6.41	note
6.45	organism
6.47	phenotype
6.49	pop_variant
6.50	product
6.66	serotype
6.67	serovar
6.68	sex
6.69	standard_name
6.70	strain
6.71	sub_clone
6.72	sub_species
6.73	sub_strain
6.75	tissue_lib
6.76	tissue_type
6.81	variety

	定義 必須 qualifier 値のフォーマット 例 コメント	特定の遺伝子の対立遺伝子に付した名称 フリーテキスト <INSDQualifier_value>adh1-1</INSDQualifier_value> ある遺伝子に存在する、遺伝子関連の feature (エクソン、コード領域等) にはすべて、同じ qualifier 「allele」の qualifier 値を用いるべきである； qualifier「allele」の qualifier 値は、定義上 qualifier 「gene」の qualifier の値と同じであってはならない； feature key「variation」と共に用いられる場合、 qualifier「allele」の qualifier 値は、その変異体の値であるべきである。
6.2.	qualifier 定義 必須 qualifier 値のフォーマット 例 コメント	anticodon tRNA のアンチコドンの位置とそれがコードするアミノ酸の位置 pos:<location>,aa:<amino_acid>,seq:<text> <location>は、アンチコドンの位置 <amino_acid>は、tRNA がコードするアミノ酸の 3 文字の略語 <text>は、アンチコドンの配列 <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>
6.3.	qualifier 定義 必須 qualifier 値のフォーマット 例 コメント	bound_moiety 特定の feature と結合する可能性のある分子/複合体の名称 フリーテキスト 言語依存：この値は、国内／広域的な手続きのために翻訳が必要な場合がある。 <INSDQualifier_value>GAL4</INSDQualifier_value> 「misc_binding」、「oriT」、「protein_bind」の各 feature には、1 つの qualifier「bound_moiety」の使用が可能。
6.4.	qualifier 定義 必須 qualifier 値のフォーマット 例	cell_line 配列が得られた細胞株 フリーテキスト <INSDQualifier_value>MCF7</INSDQualifier_value>
6.5.	qualifier 定義 必須 qualifier 値のフォーマット 例	cell_type 配列が得られた細胞の種類 フリーテキスト 言語依存：この値は、国内／広域的な手続きのために翻訳が必要な場合がある。 <INSDQualifier_value>leukocyte</INSDQualifier_value>
6.6.	qualifier 定義 必須 qualifier 値のフォーマット 例	chromosome 配列が得られた染色体 (例. 染色体番号) フリーテキスト <INSDQualifier_value>1</INSDQualifier_value> <INSDQualifier_value>X</INSDQualifier_value>
6.7	qualifier 定義 qualifier 値のフォーマット コメント	circular_RNA 例えば、遺伝子の下流のエクソンが RNA 産物の上流のエクソンの 5'に位置している場合など、このスプライスされた RNA 産物がバックプライシングによって生じた環状 RNA (circular RNA) であるため、エクソンの順番がずれていたり、重なっていたりすることを示している。 なしおよび CDS、mRNA、tRNA など、バックプライシングの結果として生成される feature に使用する必要がある。この qualifier は、スプライスイベントが 演算子「join」で示されている場合にのみ使用する必要がある (例：join(complement(69611..69724),139856..140087))
6.8.	qualifier 定義 必須 qualifier 値のフォーマット	clone および 配列が得られたクローン フリーテキスト 言語依存：この値は、国内のおよび広域的な手続き key のために翻訳が必要な場合がある。

例 コメント		<INSDQualifier_value>lambda-hIL7.3</INSDQualifier_value> ソース(source) feature には、複数の qualifier「clone」を用いてはならない；複数のクローンから得られた配列の場合は、feature key「misc_feature」およびその qualifier「note」を用いて feature table に記載しても良い。
6.9. qualifier 定義 必須 qualifier 値のフ ォーマット 例		clone_lib 配列が得られたクローンライブラリ フリーテキスト 言語依存：この値は、国内／広域的な手続きのために翻訳が必要な場合がある。 <INSDQualifier_value>lambda-hIL7</INSDQualifier_value>
6.10. qualifier 定義 必須 qualifier 値のフ ォーマット 例		codon_start その feature の最初の塩基に呼応して、コーディングされた feature の最初の完全なコドンが確認できるオフセットを示している。 1 または 2 または 3 <INSDQualifier_value>2</INSDQualifier_value>
6.11. qualifier 定義 必須 qualifier 値のフ ォーマット 例		collected_by 標本個体を採取した人の名前や機関の名称 フリーテキスト 言語依存：この値は、国内／広域的な手続きのために翻訳が必要な場合がある。 <INSDQualifier_value>Dan Janzen</INSDQualifier_value>
6.12. qualifier 定義 必須 qualifier 値のフ ォーマット 例 コメント		collection_date 標本個体が採取された日付 YYYY-MM-DD、YYYY-MM または、YYYY のみ <INSDQualifier_value>1952-10-21</INSDQualifier_value> <INSDQualifier_value>1952-10</INSDQualifier_value> <INSDQualifier_value>1952</INSDQualifier_value> 「YYYY」は暦年を表す 4 桁の数字、「MM」は暦月を表す 2 桁の数字、「DD」はある暦月の暦日を示す 2 桁の数字
6.13. qualifier 定義 必須 qualifier 値のフ ォーマット 例 コメント		compare 比較対象となる、公開されている国際塩基配列データベースの参照情報 [accession-number.sequence-version] <INSDQualifier_value>AJ634337.1</INSDQualifier_value> この qualifier は、「misc_difference」「unsure」および「variation」の各 feature で使用しても良い。1 つのフィーチャーに、内容の異なる複数の qualifier「compare」の使用が許されている。この qualifier は、一塩基多型(SNP)のような大規模な変異体に注釈を付すことを想定していない。
6.14. qualifier 定義 必須 qualifier 値のフ ォーマット 例 コメント		cultivar 配列が得られた植物の栽培品種 フリーテキスト 言語依存：この値は、国内／広域的な手続きのために翻訳が必要な場合がある。 <INSDQualifier_value>Nipponbare</INSDQualifier_value> <INSDQualifier_value>Tenuifolius</INSDQualifier_value> <INSDQualifier_value>Candy Cane</INSDQualifier_value> <INSDQualifier_value>IR36</INSDQualifier_value> qualifier「cultivar」は、人為的に選択された産物のみに適用される；自然の、名前の付いた植物や真菌類には qualifier「variety」を用いる。
6.15. qualifier 定義 必須 qualifier 値のフ ォーマット 例		dev_stage 配列が特定の発達段階にある生物から得られたものであれば、この qualifier で指定される。 フリーテキスト 言語依存：この値は、国内／広域的な手続きのために翻訳が必要な場合がある。 <INSDQualifier_value>fourth instar larva</INSDQualifier_value>

6.16.	qualifier 定義 必須 qualifier 値のフォーマット 例 コメント	direction DNA 複製の方向 left、right または both left とは(配列上で示されている)5'末端側の方向を示し、right とは 3'末端側の方向を示している。 <INSDQualifier_value>left</INSDQualifier_value> feature key「rep_origin」に qualifier「direction」を使用した場合、left、right、both の値が許容される。ただし、feature key「oriT」に qualifier「direction」を使用した場合は、left および right の値のみが許可される。 rep_origin 機能のキーに direction qualifier を使用した場合、left、right、both の値が許容される。
6.17.	qualifier 定義 必須 qualifier 値のフォーマット 例 コメント	EC_number 配列の酵素生成物の酵素コミッション番号 フリーテキスト <INSDQualifier_value>1.1.2.4</INSDQualifier_value> <INSDQualifier_value>1.1.2.-</INSDQualifier_value> <INSDQualifier_value>1.1.2.n</INSDQualifier_value> <INSDQualifier_value>1.1.2.n1</INSDQualifier_value> 国際生化学分子生物学連合の命名法委員会(NC-IUBMB)が作成したリスト中で定義されている酵素委員会番号(CE 番号)の有効値 (1992 年発表の酵素命名法勧告、アカデミックプレス、サンディエゴまたは、最新改訂版に掲載)。酵素委員会番号はピリオドで区切られた 4 つ数字の文字列を表し、文字列の最後から 3 つまでの数字は、割り当て不確かであることを示すためにダッシュ「-」で置き換えることができる。「n」を含む記号 (例: 「n」、「n1」など) は、EC 番号が割り当て待ちである場合に、数字の代わりに最後の位置に使用することができる。このような不完全な EC 番号は NC-IUBMB では承認されないため注意を要する。
6.18.	qualifier 定義 必須 qualifier 値のフォーマット 例 コメント	ecotype 遺伝学的に生育環境への適応を反映した表現型特性を示す種内集団 フリーテキスト 言語依存: この値は、国内/広域的な手続きのために翻訳が必要な場合がある。 <INSDQualifier_value>Columbia</INSDQualifier_value> このような集団の例としては、特に日当たりの良い生息環境に対応して、通常の葉よりも毛状突起が多い葉を有するようになった集団が挙げられる。「生態型」(ecotype)は、シロイヌナズナの標準的な遺伝子株に用いられることが多いが、あらゆる無菌生物に適用することができる。
6.19.	qualifier 定義 qualifier 値のフォーマット コメント	environmental_sample 大量の環境中の DNA サンプルから直接分子的に分離して得られた配列(ポリメラーゼ連鎖反応(PCR)とそれに続く産物のクローニング、変性剤濃度勾配ゲル電気泳動法(DGGE)または他の汎用手法による)で、その由来生物を確実に特定できないものを指す。環境サンプルには、臨床検体および腸管内容物、および、特定の宿主に関連する可能性のある無名の生物からの他の配列が含まれる。これらには、特定の宿主から確実に回収できる内共生体、容易に識別できるが培養されていない野外サンプルからの生物 (例. 多くのシアノバクテリア)、病害が認められる植物から確実に回収できるファイトプラズマ (純粋培養は不可能であるが) は含まれない。 なし feature key「source」と共にのみ用いられる; qualifier「environmental_sample」を用いる feature key「source」には、qualifier「isolation_source」も用いる必要があり、qualifier「environmental_sample」を用いる feature「source」には qualifier「strain」を用いてはならない。
6.20.	qualifier 定義 必須 qualifier 値のフォーマット 例	exception 標準的な生物学の法則を用いて翻訳できないコード領域であることを示している。 以下の統制語彙を用いたフレーズの 1 つを使用する。 RNA editing (ゲノム編集) rearrangement required for product (転写産物が求める再編成) annotated by transcript or proteomic data (転写データまたはプロテオミクスに関するデータを用いた詳細情報の記載) <INSDQualifier_value>RNA editing</INSDQualifier_value> <INSDQualifier_value>rearrangement required for product</INSDQualifier_value>

	コメント	RNA 演習等の生物学的メカニズムを説明するためのみに用いられる；qualifier「exception」を用いた CDS の蛋白質翻訳は、対応する概念的な翻訳とは異なる；qualifier「transl_except」を用いるのが適切である場合（例. 翻訳が終止コドンで完了する場合）、この qualifier を用いてはならない。
6.21.	qualifier 定義 必須 qualifier 値のフォーマット 例	Frequency feature の発生頻度 feature を有する個体群の比率を割合で記載するフリーテキスト <INSDQualifier_value>23/108</INSDQualifier_value> <INSDQualifier_value>1 in 12</INSDQualifier_value> <INSDQualifier_value>0.85</INSDQualifier_value>
6.22.	qualifier 定義 必須 qualifier 値のフォーマット 例 コメント	function 配列に起因する機能 フリーテキスト 言語依存：この値は、国内／広域的な手続きのために翻訳が必要な場合がある。 <INSDQualifier_value>essential for recognition of cofactor </INSDQualifier_value> function qualifier は、遺伝子名や製品名では配列に起因する機能を伝えられない場合で使用される。
6.23.	qualifier 定義 必須 qualifier 値のフォーマット 例 コメント	gene 配列領域に対応する遺伝子記号 フリーテキスト <INSDQualifier_value>ilvE</INSDQualifier_v qualifier「gene」は遺伝子記号を記載するために用いる；完全な遺伝子名を記載するには qualifier「standard_name」を用いる。
6.24.	qualifier 定義 必須 qualifier 値のフォーマット 例 コメント	gene_synonym 同義的な、置換された、廃止された、または以前の遺伝子記号 フリーテキスト 言語依存：この値は、国内／広域的な手続きのために翻訳が必要な場合がある。 <INSDQualifier_value>Hox-3.3</INSDQualifier_value> (qualifier「gene」の qualifier 値が「Hoxc6」である feature。) 遺伝子記号の動議とを示すのに役立つ場合で使用される；qualifier「gene_synonym」を用いる場合、主要な遺伝子記号は必ず qualifier「gene」で記載しなければならない。
6.25.	qualifier 定義 qualifier 値のフォーマット コメント	germline 適応免疫反応の一部として体細胞の再編成を受けていない配列；親の生殖細胞から受け継いだ再編成されていない配列である。 なし qualifier「germline」、配列の由来がガメートまたは生殖細胞であるとの情報を記載するために用いてはならない；qualifier「germline」および qualifier「rearranged」は同じ source feature 内で使用してはならない；qualifier「germline」と qualifier「rearranged」は、適応免疫反応の一部として体細胞の再編成を受ける可能性のある分子にのみ使用しなければならない；これらは、顎のある脊椎動物の t-細胞受容体(TCR)と免疫グロブリン遺伝子座、および顎のない魚（ヤツメウナギやメクラウナギ）の無関係な可変リンパ球受容体(VLR)遺伝子座である；qualifier「germline」と qualifier「rearranged」は、有頭動物(系統分類 ID = 89593)以外で使用すべきではない。
6.26.	qualifier 定義 必須 qualifier 値のフォーマット 例	haplogroup 配列多型を共有する類似したハプロタイプのグループ名称。ハプログループは、個体群移動の追跡によく用いられる。 フリーテキスト 言語依存：この値は、国内／広域的な手続きのために翻訳が必要な場合がある。 <INSDQualifier_value>H*</INSDQualifier_value>
6.27.	qualifier 定義	haplotype 同一の物理的染色体上で連鎖する対立遺伝子の組み合わせの名称。組換えがない場合、各ハプロタイプは 1 つの単位として遺伝するため、集団における遺伝子流動を追跡するのに利用されても良い。

	必須 qualifier 値のフォーマット 例	フリーテキスト <INSDQualifier_value>Dw3 B5 Cw1 A1</INSDQualifier_value>
6.28.	qualifier 定義 必須 qualifier 値のフォーマット 例	host 配列された分子が得られた生物の自然界（実験室ではなく）における宿主 フリーテキスト 言語依存：この値は、国内／広域的な手続きのために翻訳が必要な場合がある。 <INSDQualifier_value>Homo sapiens</INSDQualifier_value> <INSDQualifier_value>Homo sapiens 12 year old girl</INSDQualifier_value> <INSDQualifier_value>Rhizobium NGR234</INSDQualifier_value>
6.29.	qualifier 定義 必須 qualifier 値のフォーマット 例	identified_by 標本個体の分類学的な同定を行った専門家の名前 フリーテキスト 言語依存：この値は、国内／広域的な手続きのために翻訳が必要な場合がある。 <INSDQualifier_value>John Burns</INSDQualifier_value>
6.30.	qualifier 定義 必須 qualifier 値のフォーマット 例	isolate 配列が得られた個々の分離 フリーテキスト 言語依存：この値は、国内／広域的な手続きのために翻訳が必要な場合がある。 <INSDQualifier_value>Patient #152</INSDQualifier_value> <INSDQualifier_value>DGGE band PSBAC-13</INSDQualifier_value>
6.31.	qualifier 定義 必須 qualifier 値のフォーマット 例 コメント	isolation_source 配列が得られた生物学的サンプルの物理的、環境的および／あるいは地理学的由来を記述する。 フリーテキスト 言語依存：この値は、国内／広域的な手続きのために翻訳が必要な場合がある。 <INSDQualifier_value>rumen isolates from standard Pelleted ration-fed steer #67</INSDQualifier_value> <INSDQualifier_value>permanent Antarctic sea ice</INSDQualifier_value> <INSDQualifier_value>denitrifying activated sludge from carbon_limited continuous reactor</INSDQualifier_value> feature key「source」でのみ使用する； qualifier「environmental_sample」を用いる feature key「source」は、 qualifier「isolation_source」も用いるべきである。
6.32.	qualifier 定義 必須 qualifier 値のフォーマット 例 コメント	lab_host 配列された分子が得られた由来生物を増殖させるために実験室で使われた宿主の学名。 フリーテキスト 言語依存：この値は、国内／広域的な手続きのために翻訳が必要な場合がある。 <INSDQualifier_value>Gallus gallus</INSDQualifier_value> <INSDQualifier_value>Gallus gallus embryo</INSDQualifier_value> <INSDQualifier_value>Escherichia coli strain DH5 alpha</INSDQualifier_value> <INSDQualifier_value>Homo sapiens HeLa cells</INSDQualifier_value> 宿主生物の二名法による学名が分かっている場合は、それを使用すべきである； 宿主に関連する追加情報を記載しても良い。
6.33.	qualifier 定義 必須 qualifier 値のフォーマット 例	lat_lon 標本個体が採取された場所の地理的座標 フリーテキスト – 経緯度の座標の雛形は「d[d.dddd] N S d[dd.dddd] W E」 <INSDQualifier_value>47.94 N 28.12 W</INSDQualifier_value> <INSDQualifier_value>45.0123 S 4.1234 E</INSDQualifier_value>
6.34.	qualifier 定義 qualifier 値のフォーマット	macronuclear 示されている配列が DNA であり、大核と小核の間で染色体の分化を行う生物からのものである場合、その配列が大核 DNA からのものであることを示すために、使用する。 なし

6.35	qualifier 定義 必須 qualifier 値のフォーマット 例	map ゲノムマップ上の feature の位置 フリーテキスト <INSDQualifier_value>8q12-q13</INSDQualifier_value>
6.36.	qualifier 定義 必須 qualifier 値のフォーマット 例 コメント	mating_type 配列が得られた生物の交配タイプ；交配タイプは、原核生物および、雌雄同体の配偶子を持たずに減数分裂を行う真核生物に使用される。 フリーテキスト 言語依存：この値は、国内／広域的な手続きのために翻訳が必要な場合がある。 <INSDQualifier_value>MAT-1</INSDQualifier_value> <INSDQualifier_value>plus</INSDQualifier_value> <INSDQualifier_value>-</INSDQualifier_value> <INSDQualifier_value>odd</INSDQualifier_value> <INSDQualifier_value>even</INSDQualifier_value>" コメント qualifier「mating_type」の male と female は原核生物では有効だが、真核細胞では無効である。詳細は、qualifier「sex」のコメントを参照のこと。
6.37.	qualifier 定義 必須 qualifier 値のフォーマット 例 コメント	mobile_element_type 親 feature で記述された可動因子の識別子のタイプおよび名称 <mobile_element_type>[:<mobile_element_name>] 上記<mobile_element_type>には、以下のいずれかを用いる。 transposon retrotransposon integron insertion sequence non-LTR retrotransposon SINE MITE LINE other <INSDQualifier_value>transposon:Tnp9</INSDQualifier_value> コメント qualifier「mobile_element_type」は、feature key「mobile_element」にのみ許される。可動因子は、現在可動的な因子と、過去に可動的であった因子の両方を表すために使用されるべきである。 <mobile_element_type>の qualifier 値「other」には、<mobile_element_name>が必要である。
6.38.	qualifier 定義 必須 qualifier 値のフォーマット 例 コメント	mod_base 修飾塩基の略語 本附属書の第 2 節から、修飾塩基の略語を選択する。 <INSDQualifier_value>m5c</INSDQualifier_value> <INSDQualifier_value>OTHER</INSDQualifier_value> コメント 本附属書の第 2 節にない特定の修飾塩基は、qualifier「mod_base」の値として「OTHER」を入力し、その値として修飾塩基の完全名称と共に qualifier「note」を記載することで注釈を付す。
6.39.	qualifier 定義 必須 qualifier 値のフォーマット	mol_type 配列の分子タイプ 以下より 1 つを選択する。 genomic DNA genomic RNA mRNA tRNA rRNA other RNA other DNA transcribed RNA viral cRNA unassigned DNA unassigned RNA

例		<INSDQualifier_value>genomic DNA</INSDQualifier_value>
コメント		<INSDQualifier_value>other RNA</INSDQualifier_value> qualifier「mol_type」は feature key「source」に必須である；本 qualifier の qualifier 値「genomic DNA」は、分子が核であることを意味しない（例、細胞小器官やプラスミド DNA は、「genomic DNA」を使用して記述する必要がある。）；リボソーム RNA 遺伝子は「genomic DNA」を用いなければならない；「rRNA」は、リボソーム RNA 分子自体の配列が決定されている場合のみ適用しなければならない；「other RNA」および「other DNA」は合成分子にのみ用いなければならない；「unassigned DNA」および「unassignedRNA」は、生体内の分子が不明な場合に適用しなければならない。
6.40.	qualifier 定義 必須 qualifier 値のフォーマット	ncRNA_class parent key「ncRNA」によって記述されたノンコーディング RNA の分類の構造化された記述。 TYPE TYPE とは以下の統制語彙または統制フレーズのうちのいずれか 1 つである。 antisense_RNA autocatalytically_spliced_intron circRNA ribozyme hammerhead_ribozyme lncRNA RNase_P_RNA RNase_MRP_RNA telomerase_RNA guide_RNA sgRNA rasiRNA scRNA scaRNA siRNA pre_miRNA miRNA piRNA snoRNA snRNA SRP_RNA vault_RNA Y_RNA other
例		<INSDQualifier_value>autocatalytically_spliced_intron </INSDQualifier_value> <INSDQualifier_value>siRNA</INSDQualifier_value> <INSDQualifier_value>scRNA</INSDQualifier_value> <INSDQualifier_value>other</INSDQualifier_value>
コメント		qualifier「ncRNA_class」に用いられていない特定の ncRNA タイプは、qualifier「ncRNA_class」に「other」と入力し、qualifier「note」に新規「ncRNA_class」の簡単な説明を記述することで、注釈を付す必要がある。
6.41.	qualifier 定義 必須 qualifier 値のフォーマット 例	note コメントや追加情報 フリーテキスト 言語依存：この値は、国内／広域的な手続きのために翻訳が必要な場合がある。 <INSDQualifier_value>A comment about the feature</INSDQualifier_value>
6.42.	qualifier 定義 必須 qualifier 値のフォーマット 例 コメント	number 5'末端から 3'末端の方向に存在する遺伝因子（例、エクソンまたはイントロン）の順序を示す数字。 フリーテキスト（空白（スペース）文字は用いない） <INSDQualifier_value>4</INSDQualifier_value> <INSDQualifier_value>6B</INSDQualifier_value> 文字列には、空白文字を含まないデータ値として表される、整数、文字、または整数と文字の組み合わせに限定される。追加の用語は情報は qualifier「standard_name」で記載すべきである；例：値が「2A」である qualifier「number」や、値が「long」である qualifier「standard_name」。

6.43.	qualifier 定義 必須 qualifier 値のフォーマット 例	operon その feature が属する、単一の転写物に転写された連続した遺伝子群の名称。 フリーテキスト <INSDQualifier_value>lac</INSDQualifier_value>
6.44.	qualifier 定義 必須 qualifier 値のフォーマット 例	organelle 配列が取得された膜結合型細胞内構造の種類。 以下の統制語彙および統制句の 1 つを本 qualifier 値として用いる。 chromatophore hydrogenosome mitochondrion nucleomorph plastid mitochondrion:kinetoplast plastid:chloroplast plastid:apicoplast plastid:chromoplast plastid:cyanelle plastid:leucoplast plastid:proplastid <INSDQualifier_value>chromatophore</INSDQualifier_value> <INSDQualifier_value>hydrogenosome</INSDQualifier_value> <INSDQualifier_value>mitochondrion</INSDQualifier_value> <INSDQualifier_value>nucleomorph</INSDQualifier_value> <INSDQualifier_value>plastid</INSDQualifier_value> <INSDQualifier_value>mitochondrion:kinetoplast</INSDQualifier_value> <INSDQualifier_value>plastid:chloroplast</INSDQualifier_value> <INSDQualifier_value>plastid:apicoplast</INSDQualifier_value> <INSDQualifier_value>plastid:chromoplast</INSDQualifier_value> <INSDQualifier_value>plastid:cyanelle</INSDQualifier_value> <INSDQualifier_value>plastid:leucoplast</INSDQualifier_value> <INSDQualifier_value>plastid:proplastid</INSDQualifier_value>
6.45.	qualifier 定義 必須 qualifier 値のフォーマット 例	organism 配列された遺伝物質を記載した生物の学名（判明している場合）、またはその生物が分類されていない場合には利用可能な分類学的情報；または配列が人工的に構築された旨の表示。 フリーテキスト 言語依存：この値は、国内／広域的な手続きのために翻訳が必要な場合がある。 <INSDQualifier_value>Homo sapiens</INSDQualifier_value>
6.46.	qualifier 定義 必須 qualifier 値のフォーマット 例 コメント	PCR_primers 配列の増幅に使用された PCR プライマー。単一の qualifier「PCR_primers」は、単一の PCR 反応に使用されるすべてのプライマーを用いるべきである。単一の PCR 反応に複数のフォワードプライマーまたはリバースプライマーが存在する場合、複数の fwd_name/fwd_seq や rev_name/rev_seq の組が存在することになる。 [fwd_name: XXX1,]fwd_seq: xxxxx1,[fwd_name: XXX2,]fwd_seq: xxxxx2, [rev_name: YYY1,]rev_seq: yyyyy1,[rev_name: YYY2,]rev_seq: yyyyy2 <INSDQualifier_value>fwd_name: CO1P1, fwd_seq: ttgatttttggtcayccwgaagt,rev_name: CO1R4, rev_seq: ccwvytardcctarraartgttg</INSDQualifier_value> <INSDQualifier_value>fwd_name: hoge1, fwd_seq: cgkgtgtatcttact, rev_name: hoge2, rev_seq: cg<i>gtgtatcttact</INSDQualifier_value> <INSDQualifier_value>fwd_name: CO1P1, fwd_seq: ttgatttttggtcayccwgaagt, fwd_name: CO1P2, fwd_seq: gatacacaggtcayccwgaagt, rev_name: CO1R4, rev_seq: ccwvytardcctarraartgttg</INSDQualifier_value> fwd_seq と rev_seq はともに必須である； fwd_name と rev_name はともに任意である。両配列は 5' 末端から 3' 末端の方向に示さなければならない。配列は本附属書の第 1 節に記載されている記号を用いなければならない、修飾塩基を除いて、山かっこ(すなわち「<」または「>」)の中に記号を記載しなければならない

		ない。XML 形式では、山かっこ(「<」および「>」)は、XML の予約文字であるため、「<」と「>」に置き換えなければならない。
6.47.	qualifier 定義 必須 qualifier 値のフォーマット 例	phenotype feature によって与えられる表現型（ここで表現型とは、物理的、生化学的、行動的な特性または一連の特性と定義される）。 フリーテキスト 言語依存：この値は、国内／広域的な手続きのために翻訳が必要な場合がある。 <INSDQualifier_value>erythromycin resistance</INSDQualifier_value>
6.48.	qualifier 定義 必須 qualifier 値のフォーマット 例	plasmid 配列が得られた天然のプラスミドの名称。プラスミドは、qualifier「chromosome」または qualifier「segment」では表現できない、独立して複製する遺伝学的単位と定義される。 フリーテキスト <INSDQualifier_value>pC589</INSDQualifier_value>
6.49.	qualifier 必須 qualifier 値のフォーマット 例	pop_variant 配列の由来となったサンプルの亜集団または表現型の名称。 フリーテキスト 言語依存：この値は、国内／広域的な手続きのために翻訳が必要な場合がある。 <INSDQualifier_value>pop1</INSDQualifier_value> <INSDQualifier_value>Bear Paw</INSDQualifier_value>
6.50.	qualifier 定義 必須 qualifier 値のフォーマット 例	product feature に関連する生成物の名称。例、feature「mRNA」の mRNA、CDS の polypeptide、mat_peptide の mature peptide など。 フリーテキスト 言語依存：この値は、国内／広域的な手続きのために翻訳が必要な場合がある。 <INSDQualifier_value>trypsinogen</INSDQualifier_value> (when qualifier appears in CDS feature) <INSDQualifier_value>trypsin</INSDQualifier_value> (when qualifier appears in mat_peptide feature) <INSDQualifier_value>XYZ neural-specific transcript</INSDQualifier_value> (when qualifier appears in mRNA feature)
6.51.	qualifier 定義 必須 qualifier 値のフォーマット 例	protein_id 蛋白質配列番号（対応する feature key「CDS」およびその translation qualifier で特定されるコード配列によってコードされた蛋白質の配列を指定するために、配列表で使用される整数）。 ゼロより大きい整数 <INSDQualifier_value>89</INSDQualifier_value>
6.52.	qualifier 定義 qualifier 値のフォーマット	proviral 本 qualifier は、他の生物のゲノムに組み込まれたウイルスやファージから得られた配列にフラグを立てるために使用する。 なし
6.53.	qualifier 定義 qualifier 値のフォーマット コメント	pseudo feature key によって特定されている機能を有していたが、その機能を失った feature であることを示す qualifier。 なし qualifier「pseudo」は、以前は偽遺伝子と表現されていなかったが現在は機能を失った遺伝子であることを示すために用いられるべきである。例として、偽遺伝子化の事象発生以外の理由で翻訳機能を失ったコード領域。他の理由としては、配列決定や構築の誤りなどが考えられる。偽遺伝子に注釈を付すためには、偽遺伝子のタイプを示す qualifier「pseudogene」を使用しなければならない。
6.54.	qualifier 定義	pseudogene この feature が、feature key で指定された要素の偽遺伝子であることを示す。

必須 qualifier 値のフォーマット	<p>TYPE</p> <p>上記の TYPE は、以下の統制語彙または統制句のうちの 1 つである。</p> <p>processed</p> <p>unprocessed</p> <p>unitary</p> <p>allelic</p> <p>unknown</p>	
例	<pre><INSDQualifier_value>processed</INSDQualifier_value> <INSDQualifier_value>unprocessed</INSDQualifier_value> <INSDQualifier_value>unitary</INSDQualifier_value> <INSDQualifier_value>allelic</INSDQualifier_value> <INSDQualifier_value>unknown</INSDQualifier_value></pre>	
コメント	<p>本 qualifier 値「TYPE」の定義は以下の通り。:</p> <p>処理済み(processed) – 偽遺伝子が、mRNA が cDNA に逆転写された後、ゲノムに再統合されて生じたもの。そのため、イントロンやエクソンの構造は失われており、疑似ポリ(A)尾部を有している場合がある。</p> <p>未処理(unprocessed) – 偽遺伝子が、親遺伝子のコピーから、複製とその後の変則的な突然変異の蓄積によって生じたもの。機能的相同性と比較して、挿入、欠失、未成熟終止コドン、フレームシフトおよび同義置換の逆である非同義置換の割合増加等の変化が見られる。</p> <p>単一(unitary) – 親遺伝子が存在しない偽遺伝子。偽遺伝子が親遺伝子であり、一部の生物種では機能するが、別の生物種または系統ではなんらかの方法（インデル、変異、組み換え）で破壊されている元の遺伝子。</p> <p>対立遺伝子(allelic) – ある個体群に常に認められる遺伝子であるが、その個体群は機能的に異なる対立遺伝子も有している。すなわち、ある系統では機能が備わっている遺伝子を有しているが、別の系統は偽遺伝子を有しているという場合もある。主要組織適合抗原複合体(MHC)ハプロタイプには、対立遺伝子の偽遺伝子が存在する。</p> <p>不明(unknown) - 出願人は、偽遺伝子化の方法を知らない。</p>	
6.55.	<p>qualifier 定義</p> <p>qualifier 値のフォーマット</p> <p>コメント</p>	<p>rearranged</p> <p>イントリで記載された配列が、適応免疫反応の一部として体細胞で再配列されたものであり、親の生殖細胞から受け継いだ再配列されていない配列ではないこと。</p> <p>なし</p> <p>qualifier「rearranged」は、適応免疫反応に関与しない染色体再配列の注釈に使用してはならない； qualifier「germline」および qualifier「rearranged」は、同じソース(source) feature で使用してはならない。 qualifier「germline」および qualifier「rearranged」は、適応免疫反応の一部として体細胞再配列を起こす可能性のある分子にのみ使用しなければならない。これらは、顎のある脊椎動物の t-細胞受容体（TCR） および免疫グロブリン遺伝子座、および顎のない魚（ヤツメウナギとメクラウナギ）の無関係な可変リンパ球受容体（VLR） 遺伝子座である； qualifier「germline」および qualifier「rearranged」は有頭動物(系統分類 ID = 89593)以外では使用してはならない。</p>
6.56.	<p>qualifier 定義</p> <p>必須 qualifier 値のフォーマット</p>	<p>recombination_class</p> <p>配列内の組換えホットスポット領域の分類を構造化して記述したもの。</p> <p>TYPE</p> <p>上記の TYPE は以下の統制語彙または統制句の 1 つである。</p> <p>meiotic</p> <p>mitotic</p> <p>non_allelic_homologous</p> <p>chromosome_breakpoint</p> <p>other</p> <p>例</p> <pre><INSDQualifier_value>meiotic</INSDQualifier_value> <INSDQualifier_value>chromosome_breakpoint</INSDQualifier_value></pre> <p>コメント</p> <p>統制語彙 recombination_class に用いられていない特定の組換えクラスは、 qualifier 「recombination_class」の qualifier 値として「other」を入力し、新規の組換えクラスに関する簡単な説明を qualifier「note」に記述することにより、注釈を付さなければならない。</p>

6.57.	<p>qualifier 定義 必須 qualifier 値のフォーマット</p> <p>例</p> <p>コメント</p>	<p>regulatory_class 配列内の転写、翻訳、複製、染色体構造に関連する調節エレメントの分類を構造的に記述したもの。 TYPE 上記の TYPE は以下の統制語彙または統制句の 1 つである。 attenuator CAAT_signal DNase_I_hypersensitive_site enhancer enhancer_blocking_element GC_signal imprinting_control_region insulator locus_control_region matrix_attachment_region minus_35_signal minus_10_signal polyA_signal_sequence promoter recoding_stimulatory_region recombination_enhancer replication_regulatory_region response_element ribosome_binding_site riboswitch silencer TATA_box terminator transcriptional_cis_regulatory_region uORF other</p> <p><INSDQualifier_value>promoter</INSDQualifier_value> <INSDQualifier_value>enhancer</INSDQualifier_value> <INSDQualifier_value>ribosome_binding_site</INSDQualifier_value></p> <p>統制語彙 regulatory_class にまだ記載されていない特定の規制クラスは、qualifier 「regulatory_class」の値として「other」を入力し、新規の規制クラスについての簡単な説明を qualifier 「note」に記述することで注釈を付さなければならない。</p>
6.58.	<p>qualifier 定義</p> <p>必須 qualifier 値のフォーマット</p> <p>例</p>	<p>replace feature の location で特定された配列が、qualifier「replace」の qualifier 値で示される配列に置き換えられることを示す；Qualifier に配列が記載されていない（すなわち、qualifier 値が記載されていない）場合、欠失を意味する。 フリーテキスト</p> <p><INSDQualifier_value>a</INSDQualifier_value> <INSDQualifier_value></INSDQualifier_value> - 欠失の場合</p>
6.59.	<p>qualifier 定義</p> <p>qualifier 値のフォーマット</p> <p>コメント</p>	<p>ribosomal_slippage 蛋白質の翻訳中に、特定の配列がリボソームにプログラムされ、リボソームスリップと呼ばれるメカニズムにより、代替のリーディングフレームに変更されることがある。 なし</p> <p>feature 「CDS」の位置には、例えば[join(486..1784,1787..4810)]のような存在位置演算子を使用して、ribosomal_slippage の location を示す必要がある。</p>
6.60.	<p>qualifier 定義 必須 qualifier 値のフォーマット</p> <p>例</p>	<p>rpt_family 反復配列の種類、例えば制限酵素アルスロバクター・ルテウス(Alu)またはクレブシエラ・ニューモニエ(Kpn)。 フリーテキスト</p> <p><INSDQualifier_value>Alu</INSDQualifier_value></p>

6.61.	qualifier 定義 必須 qualifier 値のフォーマット	rpt_type 反復配列の構造と分布 以下の統制語彙または統制句の 1 つを qualifier 値に用いる。 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 例 <code><INSDQualifier_value>inverted</INSDQualifier_value></code> <code><INSDQualifier_value>long_terminal_repeat</INSDQualifier_value></code> コメント 上述の qualifier 値の定義は以下の通りである。 直列(tandem) - 同一の配列が隣り合って同じ方向で繰り返し配置された配列。 直接(direct) - 同一の配列が必ずしも隣り合って存在しないが、同じ方向で繰り返し配置された配列。 反転(inverted) - 1 つの分子上で同一の配列が逆方向に繰り返し配置された配列。 隣接(flanking) - 機能上の重要な影響を及ぼす配列の外に存在する反復配列(例 トランスポゾンの挿入部位) 入れ子(nested) - 他の要素の挿入により中断されているが、繰り返し配置された配列。 分散(dispersed) - 遺伝子全体に散在している反復配列。 末端(terminal) - 機能上の重要な影響を及ぼす配列上の末端で、繰り返し配置された配列 (例 末端に長い反復配列を有する(LTR 型)トランスポゾン) long_terminal_repeat - 定義された配列の両端で、同じ方向に繰り返し配置された配列。一般的にレトロウイルスの配列上で認められる種類の反復配列。 non_ltr_retrotransposon_polymeric_tract - 非 LTR 型レトロトランスポゾンの配列上のポリ(dA)等の、ポリトラクト。 centromeric_repeat - セントロメアの反復配列の中で認められる、反復配列。 telomeric_repeat - テロメアの配列の中で認められる、反復配列。 x_element_combinatorial_repeat - X エlementとテロメア、またはその隣接する Y'Elementの間に存在する反復配列。 y_prime_element - 1 コピーまたは 2 から 4 コピーの縦列反復であるかいずれか一方の、テロメア反復配列または、X Elementのコンビナトリアル反復配列に隣接して配置された反復配列の領域。 その他(other) - 他の値では表現できない重要な属性を示す反復配列。
6.62.	qualifier 定義 必須 qualifier 値のフォーマット	rpt_unit_range 範囲で表現された反復単位の location。 <code><base_range></code>

		<p>上述の<base_range>には、反復ユニットの先頭の塩基の位置と末端の塩基の位置が2つのドットで区切られて示されている。</p> <p>例 コメント <INSDQualifier_value>202..245</INSDQualifier_value> feature key「oriT」および「repeat_reagon」で指定された領域内で、反復単位を構成する配列の塩基の範囲を示すために使用する。</p>
6.63.	<p>qualifier 定義 必須 qualifier 値のフォーマット 例 コメント</p>	<p>rpt_unit_seq 反復配列の同一性。 フリーテキスト</p> <p><INSDQualifier_value>aagggc</INSDQualifier_value> <INSDQualifier_value>ag(5)tg(8)</INSDQualifier_value> <INSDQualifier_value>(AAAGA)6(AAAA)1(AAAGA)12</INSDQualifier_value> feature key「oriT」および「repeat_region」で指定された領域内の反復単位を構成する文字通りの配列を示すために使用する。</p>
6.64	<p>qualifier 定義 必須 qualifier 値のフォーマット 例 コメント</p>	<p>satellite サテライト DNA マーカーの識別子であり、短い基本反復単位の多数のタンデム反復（同一または関連）で構成される。</p> <p><satellite_type>[:<class>][<identifier>] 上記の<satellite_type>は以下のいずれかの1つである。 satellite; microsatellite; minisatellite</p> <p>例 コメント <INSDQualifier_value>satellite: S1a</INSDQualifier_value> <INSDQualifier_value>satellite: alpha</INSDQualifier_value> <INSDQualifier_value>satellite: gamma III</INSDQualifier_value> <INSDQualifier_value>microsatellite: DC130</INSDQualifier_value> 多くのサテライトは、塩基組成やその他の特性が他のゲノムと異なるため、それらを識別することができる。</p>
6.65.	<p>qualifier 定義 必須 qualifier 値のフォーマット 例</p>	<p>segment 配列されたウイルスまたはファージのセグメントの名称。 フリーテキスト</p> <p>例 <INSDQualifier_value>6</INSDQualifier_value></p>
6.66.	<p>qualifier 定義 必須 qualifier 値のフォーマット 例 コメント</p>	<p>serotype 抗原性を特徴とする種の血清学的多様性。 フリーテキスト 言語依存：この値は、国内／広域的な手続きのために翻訳が必要な場合がある。</p> <p>例 コメント <INSDQualifier_value>B1</INSDQualifier_value> feature key「source」と共にのみ用いられる；国際原核生物命名規約(bacteriological Code)では、原核生物には「血清型」ではなく「血液型亜型」という用語を使用することを推奨している；国際原核生物命名規約(1990年版)、附属書 10.8、「種内用語」を参照のこと。</p>
6.67.	<p>qualifier 定義 必須 qualifier 値のフォーマット 例 コメント</p>	<p>serovar 抗原性を特徴とする種（通常は原核生物）の血清学的変異体 フリーテキスト 言語依存：この値は、国内／広域的な手続きのために翻訳が必要な場合がある。</p> <p>例 コメント <INSDQualifier_value>O157:H7</INSDQualifier_value> feature key「source」と共にのみ用いられる；国際原核生物命名規約では、原核生物に対して「serotype」の代わりに「serovar」という用語を使用することを推奨している。</p>
6.68.	<p>qualifier 定義 必須 qualifier 値のフォーマット 例</p>	<p>sex 配列が得られた生物の性別；真核生物で減数分裂を行い、性的に異なる配偶子を持つ場合は性別が用いられる。 フリーテキスト 言語依存：この値は、国内／広域的な手続きのために翻訳が必要な場合がある。</p> <p>例 <INSDQualifier_value>female</INSDQualifier_value> <INSDQualifier_value>male</INSDQualifier_value></p>

		<p><INSDQualifier_value>hermaphrodite</INSDQualifier_value> <INSDQualifier_value>unisexual</INSDQualifier_value> <INSDQualifier_value>bisexual</INSDQualifier_value> <INSDQualifier_value>asexual</INSDQualifier_value> <INSDQualifier_value>monoecious</INSDQualifier_value> [or monecious] <INSDQualifier_value>dioecious</INSDQualifier_value> [or diecious] qualifier「sex」 (qualifier「mating_type」ではなく)は、後生動物、有胚植物類、紅藻類 & 褐藻類に 関して用いるべきである；細菌、古細菌&菌類界に関しては、qualifier「mating_type」 (qualifier 「sex」ではなく)を用いるべきである；ウイルスに関しては、qualifier「sex」および qualifier「mating_type」 は用いるべきではない；その他(上述の分類群以外)の詳細情報を記載する場合で、上の例で用いられて いる用語を qualifier「sex」の qualifier 値として用いない場合は、qualifier「mating_type」を用いるべき である。</p>
6.69.	qualifier 定義 必須 qualifier 値のフ ォーマット 例 コメント	standard_name この feature の標準的な名称。 フリーテキスト 言語依存：この値は、国内／広域的な手続きのために翻訳が必要な場合がある。 <INSDQualifier_value>dotted</INSDQualifier_value> 遺伝子名を省略せずに記載する場合は qualifier「standard_name」を使用し、qualifier「gene」を用い て遺伝子コードを与える (例えば、上の例では、qualifier「gene」の値は「Dt」)。
6.70.	qualifier 定義 必須 qualifier 値のフ ォーマット 例 コメント	strain 配列が得られた strain フリーテキスト 言語依存：この値は、国内／広域的な手続きのために翻訳が必要な場合がある。 <INSDQualifier_value>BALB/c</INSDQualifier_value> qualifier「strain」を記載する feature エントリには、qualifier「environmental_sample」を用いてはなら ない。
6.71.	qualifier 定義 必須 qualifier 値のフ ォーマット 例 コメント	sub_clone 配列が得られたサブクローン フリーテキスト 言語依存：この値は、国内／広域的な手続きのために翻訳が必要な場合がある。 <INSDQualifier_value>lambda-hIL7.20g</INSDQualifier_value> ソース(source) feature は、2 つ以上の qualifier「sub_clone」を用いてはならない；複数のサブクローン から得られた配列であることを示すために、複数のソース(source)を feature key「misc_feature」およびそ の qualifier「note」を用いて詳細情報を記載しても良い。
6.72.	qualifier 定義 必須 qualifier 値のフ ォーマット 例	sub_species 配列が得られた生物の亜種の名称。 フリーテキスト 言語依存：この値は、国内／広域的な手続きのために翻訳が必要な場合がある。 <INSDQualifier_value>lactis</INSDQualifier_value>
6.73.	qualifier 定義 必須 qualifier 値のフ ォーマット 例 コメント	sub_strain 親 strain (qualifier「strain」に注釈を付すべき) から派生した、配列が得られた遺伝子的またはその他 の方法で改変された sub_strain の名称または識別子。 フリーテキスト 言語依存：この値は、国内／広域的な手続きのために翻訳が必要な場合がある。 <INSDQualifier_value>abis</INSDQualifier_value> ソース(source) feature の qualifier「strain」と共に用いられなければならない；親の系統が与えられてい ない場合は、qualifier「sub_strain」ではなく qualifier「strain」を用いて亜系統に注釈を付すべきである。 例えば、qualifier「strain」 (qualifier 値は K-12)、qualifier「sub_strain」(qualifier 値は MG1655)ま たは、qualifier「strain」 (qualifier 値は MG1655)のいずれかである。
6.74.	qualifier 定義 必須 qualifier 値のフ ォーマット	tag_peptide tmRNA のタンパク質分解タグ用ポリペプチドをコードする塩基位置とその終止コドン。 <base_range> 上述の<base_range>には、蛋白質分解のタグが存在する配列上の先頭の塩基の位置と末端の塩基 の位置が 2 つのドットで区切られて示されている。

例 コメント		<INSDQualifier_value>90..122</INSDQualifier_value> タグ ペプチドに対応するアミノ酸配列は、5'末端の部分的な feature「CDS」を示すものとして注釈を付すべきことを推奨する；例： <90...122> の location にある CDS。
6.75. qualifier 定義 必須 qualifier 値のフォーマット 例	tissue_lib 配列が得られた組織ライブラリ。 フリーテキスト 言語依存：この値は、国内／広域的な手続きのために翻訳が必要な場合がある。	<INSDQualifier_value>tissue library 772</INSDQualifier_value>
6.76. qualifier 定義 必須 qualifier 値のフォーマット 例	tissue_type 配列が得られた組織のタイプ フリーテキスト 言語依存：この値は、国内／広域的な手続きのために翻訳が必要な場合がある。	<INSDQualifier_value>liver</INSDQualifier_value>
6.77. qualifier 定義 必須 qualifier 値のフォーマット 例 コメント	transl_except 翻訳上の例外；ある 1 つのコドンの翻訳が、qualifier「organism」または qualifier「transl_table」に記載されている遺伝コードと一致しない状態。 (pos: <location,aa>:<amino_acid>) ここで、<amino_acid>は、base_range の位置にあるコドンによってコードされるアミノ酸の 3 文字の略語である。	<INSDQualifier_value>(pos:213..215,aa:Trp) </INSDQualifier_value> <INSDQualifier_value>(pos:462..464,aa:OTHER) </INSDQualifier_value> <INSDQualifier_value>(pos:1017,aa:TERM) </INSDQualifier_value> <INSDQualifier_value>(pos:2000..2001,aa:TERM) </INSDQualifier_value> アミノ酸が本附属書の第 3 節に記載されている特定のアミノ酸でない場合、「OTHER」を <amino_acid>として使用し、例外的なアミノ酸の名称を qualifier「note」に記載する；修飾アミノ酸であるセレンシステインについては、<amino_acid>に 3 文字の略語「Sec」（アミノ酸配列内の 1 文字の記号「U」）を使用し、修飾アミノ酸であるピロリジンについては、<amino_acid>に 3 文字の略語「Pyl」（アミノ酸配列内の 1 文字の記号「O」）を使用する；TAA 終止コドンが mRNA に 3'A 残基を付加することで完成する部分終止コドンの場合、その location には単一の base_position または base_range のいずれかを使用し、「mRNA に 3'A 残基を付加することで完成する終止コドン」を示す qualifier「note」と併せて用いる（上記の 3 番目と 4 番目の例を参照されたい）。
6.78. qualifier 定義 必須 qualifier 値のフォーマット 例 コメント	transl_table 普遍的または標準的な遺伝コード表以外の場合に使用される遺伝コード表の定義。使用されるテーブルは本附属書に記載がある。 <integer> 上記の<integer>とは、遺伝コード表に割り振られた番号である。	<INSDQualifier_value>3</INSDQualifier_value> 上記の例では、酵母のミトコンドリアの遺伝コード表を用いて配列表を作成している。 feature key「CDS」の詳細情報が qualifier「transl_table」に記載されていない場合は、CDS は標準コード（すなわち Universal Genetic Code）を使用して翻訳されます。指定されたテーブルの範囲外の遺伝子コードの例外は qualifier「transl_except」で報告される。
6.79. qualifier 定義 qualifier 値のフォーマット コメント	trans_splicing 2 つの RNA 分子のエクソンが連結し、成熟 RNA を形成することを示している。 なし	トランスプライシングの結果産生された feature（例えば CDS、mRNA 等）の詳細情報を記載するために本 qualifier を用いるべきである。またこの qualifier は、スプライシングの発生が feature location の演算子「joint」、例 joint (complement (69611..69724),139856..140087)を用いて示されている場合にのみ使用されなければならない。
6.80. qualifier 定義	translation 標準（またはユニバーサル）遺伝コード、または qualifier「transl_table」で指定された遺伝コード表や qualifier「transl_except」の例外によって決められた遺伝コード表のいずれかに準じた 1 文字の略号によるアミノ酸配列。	

必須 qualifier 値のフォーマット 例 コメント	本附属書の第 3 節に記載されている 1 文字のアミノ酸の記号を用いて記載した、連続した 1 本のアミノ酸配列。記号「X」は AA 等の例外的なアミノ酸に用いる記号である。 <INSDQualifier_value>MASTFPPWYRGCSTPSLKGLIMCTW</INSDQualifier_value> CDS フィーチャーと共にのみ使用される；翻訳産物に 4 つ以上の特定のアミノ酸が含まれる場合、qualifier「protein_id」を付さなければならない；遺伝コード表の定義および location について transl_table を参照されたい；feature「CDS」を追加で注釈を付すには、translation qualifier、「pseudo」、「pseudogene」のいずれかのみが許容されている。
<hr/>	
6.81. qualifier 定義	variety 配列の元となった生物の種類（= 変種、正式なリンネ式分類法の階級）
必須 qualifier 値のフォーマット 例 コメント	フリーテキスト 言語依存：この値は、国内／広域的な手続きのために翻訳が必要な場合がある。 <INSDQualifier_value>insularis</INSDQualifier_value> 植物の栽培品種（すなわち人為的な選択により栽培された植物）を記載するには qualifier「cultivar」を用いる；植物および菌類の変種以外の変種は、qualifier「note」を用いて注釈を付すべきであり、その際、qualifier 値として例えば次を用いる； <INSDQualifier_value>breed:Cukorova</INSDQualifier_value>。

第7節：アミノ酸配列の feature keys

ここでは、アミノ酸配列に使用できる feature key の一覧を示している。feature key はアルファベット順に並んでいる。

7.1.	feature key 定義 任意の qualifier コメント	ACT_SITE 酵素の活性に関与するアミノ酸 qualifier「note」 活性部位に存在するアミノ酸残基は、feature key「ACT_SITE」を用いて個別に注釈を付さなければならない。また対応するアミノ酸残基の番号を、feature location の要素に存在位置記述子として記載しなければならない。
7.2.	feature key 定義 必須 qualifier コメント	BINDING 任意の化学基(補酵素、補欠分子族等)の結合部位。基の化学的な性質は qualifier「note」を用いて記載する。 qualifier「note」 qualifier「note」の値の例としては、「Heme(covalent)」および「Chloride」がある。必要に応じて feature key「BINDING」ではなく、feature key「CA_BIND」、feature key「DNA_BIND」、feature key「METAL」および feature key「NP_BIND」を用いるべきである。
7.3.	feature key 定義 任意の qualifier	CA_BIND カルシウム結合の領域の範囲 qualifier「note」
7.4.	feature key 定義 必須 qualifier コメント	CARBOHYD グリコシル化した部位 qualifier「note」 本 key は、糖類（単糖または多糖）が蛋白質の残基に付加していることを示している。この様な種類の蛋白質への付加(C-, N- または O-結合グリコシル化)は、qualifier「note」に記載する。還元末端糖の性質が分かっている場合は、その略語を括弧の間に記載する。還元末端糖の略語の後に3つのドット「...」が続く場合は、糖鎖がさらに続いていることを示している。それとは逆に、ドットが1つも付いていない場合は単糖類が結合していることを意味している。qualifier「note」の qualifier 値の例には、次がある；N-結合(GlcNAc...)；O-結合(GlcNAc)；O-結合(Glc...)；C-結合(Man)部分；O-結合(Ara...)。
7.5.	feature key 定義 任意の qualifier	CHAIN 成熟蛋白質のポリペプチド鎖の長さ qualifier「note」
7.6.	feature key 定義 任意の qualifier	COILED 二重コイル構造を有する領域の範囲 qualifier「note」
7.7.	feature key 定義 任意の qualifier	COMPBIAS compositionally biased region(CBR。組成の偏った領域) qualifier「note」
7.8.	feature key 定義 任意の qualifier コメント	CONFLICT 異なるソースが異なる配列を報告すること qualifier「note」 qualifier「note」の qualifier 値の例；Missing; K -> Q; GSDSE -> RIRLR; V -> A.
7.9.	feature key 定義 必須 qualifier コメント	CROSSLNK 翻訳語に形成されたアミノ酸同士の結合 qualifier「note」 2つの蛋白質の間（鎖間架橋）、または同じ蛋白質の2つの部分の間（鎖内架橋）に形成される様々なタイプの共有結合。但しジスフィド結合により形成される架橋には、feature key「DISULFID」が使用される。鎖間架橋の場合、feature location 要素の存在位置記述子は、他の蛋白質に架橋されたアミノ酸の残基番号となる。鎖内架橋の場合、feature location 要素の存在位置記述子は、架橋されたアミノ酸の残基番号を「x..y」形式で表したもので、例えば、「42...50」のようになる。qualifier「note」は、架橋の性質を示すもので、少なくとも結合体の名前と関係する2つのアミノ酸の同一性を明示する。

		qualifier「note」の値の例。「イソグルタミルシステインチオエステル (Cys-Gln)」、「β-メチルランチオニン (Cys-Thr)」、「グリシルリジンイソペプチド (Lys-Gly) (コヒキチンの G-Cter との鎖間)」。
7.10.	feature key 定義 必須 qualifier コメント	DISULFID ジスフィド結合 qualifier「note」 鎖間ジスフィド結合の場合、feature location の要素の存在位置記述子は、他の蛋白質に結合しているシステインの残基番号となる。鎖内架橋の場合、feature location の要素の存在位置記述子は、リンクされたシステインの残基番号を「x..y」形式で表したもので、例えば「42..50」ようになる。鎖間ジスフィド結合の場合、note qualifier は、他の蛋白質を特定することで、架橋の性質を示す。
7.11.	feature key 定義 必須 qualifier コメント	DNA_BIND DNA 結合領域の範囲 qualifier「note」 DNA 結合領域の性質は、qualifier「note」を用いて記載する。その際の qualifier 値の例としては、qualifier「note」の例として、「Homeobox」および「Myb 2」がある。
7.12.	feature key 定義 必須 qualifier コメント	DOMAIN 特徴的な三次元構造または折り畳まれた構造を有する、二次構造の特殊な組み合わせとして定義されるドメインの範囲。 qualifier「note」 ドメインのタイプは qualifier「note」を用いて記載する。ドメインのコピーが複数存在する場合は番号を付与する。この qualifier「note」の qualifier 値の例としては、「Ras-GAP」および「Cadherin1」がある。
7.13.	feature key 定義 任意の qualifier コメント	HELIX 二次構造、すなわちヘリックス (例 αヘリックス、3(10)ヘリックスまたは Pi ヘリックス (TT ヘリックス)) qualifier「note」 feature key「HELIX」は三次構造が判明している蛋白質にのみ用いる。二次構造は 3 種類のみ指定されている (ヘリックス [key「HELIX」]、βストランド [key「STRAND」]およびターン [key「TURN」])。これらの種類の 1 つに該当しない残基は、「loop」または「random-coil」構造である。
7.14.	feature key 定義 任意の qualifier コメント	INIT_MET イニシエーターメチオニン qualifier「note」 feature location の要素の存在位置記述子は「1」である。本 feature key は、N 末端のメチオニンが切断されていることを示している。イニシエーターメチオニンが切断されていない場合は使用されない。
7.15.	feature key 定義 任意の qualifier	INTRAMEM 細胞膜に位置する領域の範囲で、細胞膜を横切らないもの。 qualifier「note」
7.16.	feature key 定義 必須 qualifier コメント	LIPID 脂質の部分との共有結合 qualifier「note」 共有結合している脂質の部分の化学的性質は、qualifier「note」を用いて記載するが、少なくとも脂質修飾を受けたアミノ酸の名称を記載する。qualifier「note」の値の例には、「N-myristoyl glycine」、「GPI-anchor amidated serine」、「S-diacylglycerol cysteine」など。
7.17.	feature key 定義 必須 qualifier コメント	METAL 金属イオンの結合部位 qualifier「note」 qualifier「note」は金属の性質を示す。qualifier「note」の qualifier 値の例には、「Iron (heme axial ligand)」および「Copper」がある。
7.18.	feature key 定義 必須 qualifier コメント	MOD_RES 残基の翻訳後修飾 qualifier「note」 修飾を受けた残基の化学的性質は qualifier「note」を用いて記載するが、少なくとも翻訳後修飾を受けたアミノ酸の名称を示す。修飾アミノ酸が本附属書の第 4 節に記載されている場合は、完全な名称の代わりに略語を使用しても良い。qualifier「note」の qualifier 値の例には、「3-Hyp」および「MeLys」および「N-6-methyllysine」がある。

7.19.	feature key 定義 任意の qualifier	MOTIF 生物学的に関心のある、短い配列モチーフ(最大で 20 個のアミノ酸が連続した配列) qualifier「note」
7.20.	feature key 定義 任意の qualifier	MUTAGEN 変異により実験的に変化した部位 qualifier「note」
7.21.	feature key 定義 任意の qualifier コメント	NON_STD 非標準アミノ酸 qualifier「note」 この key は、アミノ酸配列内の非標準アミノ酸であるセレノスチン(U)およびピロリシン(O)の出現率のみを表している。
7.22.	feature key 定義 任意の qualifier コメント	NON_TER 配列の端にある残基は末端の残基ではないこと。 qualifier「note」 配列の先頭に適用する場合、これは最初の位置が完全な分子の N-末端ではないことを意味している。最後の位置に適用する場合、この位置が完全な分子の C-末端ではないことを意味している。
7.23.	feature key 定義 必須 qualifier コメント	NP_BIND ヌクレオチドのリン酸結合領域の範囲 qualifier「note」 ヌクレオチドのリン酸の性質は、qualifier「note」を用いて記載する。その際の qualifier 値の例には、「ATP」および「FAD」がある。
7.24.	feature key 定義 任意の qualifier	PEPTIDE 放出された活性ポリペプチドの範囲 qualifier「note」
7.25.	feature key 定義 任意の qualifier	PROPEP プロペプチドの範囲 qualifier「note」
7.26.	feature key 定義 任意の qualifier	REGION 配列内の関心領域の範囲 qualifier「note」
7.27.	feature key 定義 任意の qualifier	REPEAT 内部配列の反復の範囲 qualifier「note」
7.28.	feature key 定義 任意の qualifier	SIGNAL シグナル配列の長さ (プレペプチド) qualifier「note」
7.29.	feature key 定義 必須 qualifier コメント	SITE 他の feature key で定義されていない、配列上の関心のある単一のアミノ酸部位。2 つの隣接するアミノ酸の位置で示される、アミノ酸の結合にも適用できる。 qualifier「note」 修飾アミノ酸の注釈に使用される場合、qualifier「note」の値は、本附属書の第 4 節で示した略語か、もしくは完全に、省略されていない修飾アミノ酸の名称のいずれかでなければならない。
7.30.	feature key 定義 必須 qualifier 任意の qualifier	source その配列の起源(source)を示す。この key は必須である。全ての配列には、配列全体にまたがる単一の feature「source」がある。 mol_type organism qualifier「note」

7.31.	featurekey 定義 任意の qualifier コメント	STRAND 2 次構造；βストランド；例えば水素結合したβストランドまたは孤立したβブリッジの残基 qualifier「note」 この key は、三次構造が判明している蛋白質にのみ使用する。二次構造は 3 種類のみ指定されている（ヘリックス構造 [key HELIX]、βストランド構造 [key STRAND]およびターン構造 [key TURN]）。これらのクラスのいずれかに該当しない残基は、「ループ」または「ランダムコイル」構造である。
7.32.	feature key 定義 任意の qualifier	TOPO_DOM トポロジカルドメイン qualifier「note」
7.33.	feature key 定義 任意の qualifier	TRANSMEM 膜貫通領域の範囲 qualifier「note」
7.34.	feature key 定義 任意の qualifier	TRANSIT トランジットペプチドの範囲（ミトコンドリア、葉緑体、チラコイド、シアネル、ペルオキシソーム等） qualifier「note」
7.35.	feature key 定義 任意の qualifier コメント	TURN 二次構造のターン。例えば、水素結合したターン（3-ターン、4-ターンまたは 5-ターン） qualifier「note」 この feature は、三次構造が判明している蛋白質にのみ使用する。二次構造は 3 種類のみ指定されている（ヘリックス構造 [key HELIX]、βストランド構造 [key STRAND]およびターン構造 [key TURN]）。これらのクラスのいずれかに該当しない残基は、「ループ」または「ランダムコイル」構造である。
7.36.	feature key 定義 任意の qualifier コメント	UNSURE 配列内の不確実な点 qualifier「note」 アミノ酸配列内で、著者が配列表示について確信が持てない領域を記述するために仕様する。
7.37.	feature key 定義 任意の qualifier	VARIANT 配列内の変異体の存在を著者が報告する qualifier「note」
7.38.	feature key 定義 任意の qualifier	VAR_SEQ 選択的スプライシング、代替プロモーターの使用、代替開始およびリボソームフレームシフトによって生じた配列変異体の説明 qualifier「note」
7.39.	feature key 定義 任意の qualifier コメント	ZN_FING ジンクフィンガー領域の範囲 qualifier「note」 ジンクフィンガーのタイプは、qualifier「note」に記載する。例えば、「GATA-type」、「NR C4-type」がある。

第 8 節：アミノ酸配列の qualifiers

本節では、アミノ酸配列に使用できる qualifier の一覧を示している。

qualifier 値のフォーマットが言語に依存するフリーテキストの場合、以下のいずれかを使用しなければならない。

- 1) 要素 INSDQualifier_value、または
- 2) 要素 NonEnglishQualifier_value、または
- 3) 要素 INSDQualifier_value と要素 NonEnglishQualifier_value の両方。

値のフォーマットが言語依存のフリーテキストとして識別されない場合、要素 INSDQualifier_value を使用しなければならず、要素 NonEnglishQualifier_value は使用してはならない。

注意：「フリー・テキスト」の値の形式で提供される qualifier 値は、国内または広域的な手続きのために翻訳が必要な場合がある。次の表に示す qualifier は、言語に依存したフリーテキストの値を持つとみなされる。

表 6: 言語に依存したフリーテキスト値を持つアミノ酸配列の qualifier の一覧

節	言語に依存したフリーテキストの値
8.2	note
8.3	organism

8.1.	<p>qualifier 定義 必須 qualifier 値のフォーマット 例 コメント</p>	<p>mol_type 配列の生体内における分子型 protein <INSDQualifier_value>protein</INSDQualifier_value> feature key source の必須 qualifier は「mol_type」。</p>
8.2.	<p>qualifier 定義 必須 qualifier 値のフォーマット 例 コメント</p>	<p>note コメントおよび追記のすべて フリーテキスト 言語依存：この値は、国内／広域的な手続きのために翻訳が必要な場合がある。 <INSDQualifier_value>Heme (covalent)</INSDQualifier_value> 以下の feature key の必須 qualifier は qualifier「note」である。 feature key「BINDING」、feature key「CARBOHYD」、feature key「CROSSLNK」、feature key「DISULFID」、feature key「DNA_BIND」、feature key「DOMAIN」、feature key「LIPID」、feature key「METAL」、feature key「MOD_RES」、feature key「NP_BIND」、feature key「SITE」および、feature key「ZN_FING」。</p>
8.3.	<p>qualifier 定義 必須 qualifier 値のフォーマット 例 コメント</p>	<p>organism ペプチドを有する生物の学名 フリーテキスト 言語依存：この値は、国内／広域的な手続きのために翻訳が必要な場合がある。 <INSDQualifier_value>Homo sapiens</INSDQualifier_value> qualifier「organism」は、feature key「source」の必須 qualifier である。</p>

21 - Trematode Mitochondrial Code	
AAs =	FFLLSSSSYY**CCWLLLLLPPPPHHQRRRRRIIMTTTTNNKSSSSVVVVAAAADDEEGGGG
Starts =	-----M-----M-----
Base1 =	ttttttttttttttttcccccccccccccccccaaaaaaaaaaaaaaaaaaagggggggggggggggg
Base2 =	tttcccccaaaaggggtttcccccaaaaggggtttcccccaaaaggggtttcccccaaaagggg
Base3 =	tcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcag
22 - Scenedesmus obliquus Mitochondrial Code	
AAs =	FFLLSS*SY*Y*LCC*WLLLLLPPPPHHQRRRRRIIMTTTTNNKSSRRVVVVAAAADDEEGGGG
Starts =	-----M-----
Base1 =	ttttttttttttttttcccccccccccccccccaaaaaaaaaaaaaaaaaaagggggggggggggggg
Base2 =	tttcccccaaaaggggtttcccccaaaaggggtttcccccaaaaggggtttcccccaaaagggg
Base3 =	tcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcag
23 - Thraustochytrium Mitochondrial Code	
AAs =	FF*LSSSSYY**CC*WLLLLLPPPPHHQRRRRRIIMTTTTNNKSSRRVVVVAAAADDEEGGGG
Starts =	-----M--M-----M-----
Base1 =	ttttttttttttttttcccccccccccccccccaaaaaaaaaaaaaaaaaaagggggggggggggggg
Base2 =	tttcccccaaaaggggtttcccccaaaaggggtttcccccaaaaggggtttcccccaaaagggg
Base3 =	tcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcag
24 - Pterobranchia Mitochondrial Code	
AAs =	FFLLSSSSYY**CCWLLLLLPPPPHHQRRRRRIIMTTTTNNKSSSKVVVVAAAADDEEGGGG
Starts =	---M-----M-----M-----M-----
Base1 =	ttttttttttttttttcccccccccccccccccaaaaaaaaaaaaaaaaaaagggggggggggggggg
Base2 =	tttcccccaaaaggggtttcccccaaaaggggtttcccccaaaaggggtttcccccaaaagggg
Base3 =	tcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcag
25 - Candidate Division SR1 and Gracilibacteria Code	
AAs =	FFLLSSSSYY**CCGWLLLLLPPPPHHQRRRRRIIMTTTTNNKSSRRVVVVAAAADDEEGGGG
Starts =	---M-----M-----M-----
Base1 =	ttttttttttttttttcccccccccccccccccaaaaaaaaaaaaaaaaaaagggggggggggggggg
Base2 =	tttcccccaaaaggggtttcccccaaaaggggtttcccccaaaaggggtttcccccaaaagggg
Base3 =	tcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcag
26 - Pachysolen tannophilus Nuclear Code	
AAs =	FFLLSSSSYY**CC*WLLLLAPPPPHHQQRRRRRIIMTTTTNNKSSRRVVVVAAAADDEEGGGG
Starts =	-----M-----M-----
Base1 =	ttttttttttttttttcccccccccccccccccaaaaaaaaaaaaaaaaaaagggggggggggggggg
Base2 =	tttcccccaaaaggggtttcccccaaaaggggtttcccccaaaaggggtttcccccaaaagggg
Base3 =	tcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcag
27 - Karyorelict Nuclear Code	
AAs =	FFLLSSSSYYQCCWLLLLLPPPPHHQRRRRRIIMTTTTNNKSSRRVVVVAAAADDEEGGGG
Starts =	-----*-----M-----
Base1 =	ttttttttttttttttcccccccccccccccccaaaaaaaaaaaaaaaaaaagggggggggggggggg
Base2 =	tttcccccaaaaggggtttcccccaaaaggggtttcccccaaaaggggtttcccccaaaagggg
Base3 =	tcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcag

28 - Condylostoma Nuclear Code[↗]	
AAs =	FFLLSSSSYYQCCWLLLLPPPPHHQRRRRIIIMTTTTNKKSSRRVVVVAAAADDEEGGGG..
Starts =	-----**-----M-----..
Base1 =	ttttttttttttttttcccccccccccccccccaaaaaaaaaaaaaaaaaaggggggggggggggggg..
Base2 =	ttttccccaaaaggggttttccccaaaaggggttttccccaaaaggggttttccccaaaagggg..
Base3 =	tcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcag..
..	
29 - Mesodinium Nuclear Code[↗]	
AAs =	FFLLSSSSYYYYCC*WLLLLPPPPHHQRRRRIIIMTTTTNKKSSRRVVVVAAAADDEEGGGG..
Starts =	-----M-----..
Base1 =	ttttttttttttttttcccccccccccccccccaaaaaaaaaaaaaaaaaaggggggggggggggggg..
Base2 =	ttttccccaaaaggggttttccccaaaaggggttttccccaaaaggggttttccccaaaagggg..
Base3 =	tcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcag..
..	
30 - Peritrich Nuclear Code[↗]	
AAs =	FFLLSSSSYYEECC*WLLLLPPPPHHQRRRRIIIMTTTTNKKSSRRVVVVAAAADDEEGGGG..
Starts =	-----M-----..
Base1 =	ttttttttttttttttcccccccccccccccccaaaaaaaaaaaaaaaaaaggggggggggggggggg..
Base2 =	ttttccccaaaaggggttttccccaaaaggggttttccccaaaaggggttttccccaaaagggg..
Base3 =	tcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcag..
..	
31 - Blastocrithidia Nuclear Code[↗]	
AAs =	FFLLSSSSYYEECCWLLLLPPPPHHQRRRRIIIMTTTTNKKSSRRVVVVAAAADDEEGGGG..
Starts =	-----**-----M-----..
Base1 =	ttttttttttttttttcccccccccccccccccaaaaaaaaaaaaaaaaaaggggggggggggggggg..
Base2 =	ttttccccaaaaggggttttccccaaaaggggttttccccaaaaggggttttccccaaaagggg..
Base3 =	tcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcag..
..	
33 - Cephalodiscidae Mitochondrial UAA-Tyr Code[↗]	
AAs =	FFLLSSSSYYY=CCWLLLLPPPPHHQRRRRIIIMTTTTNKKSSSKVVVVAAAADDEEGGGG..
Starts =	---M-----M-----M-----M-----..
Base1 =	ttttttttttttttttcccccccccccccccccaaaaaaaaaaaaaaaaaaggggggggggggggggg..
Base2 =	ttttccccaaaaggggttttccccaaaaggggttttccccaaaaggggttttccccaaaagggg..
Base3 =	tcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcag..
..	

[次頁から附属書Ⅱが続く]

附属書 II

配列表の文書型定義(DTD)

Version 1.3

2021年11月5日開催の第9回WIPO標準委員会(CWS)会合で採択

<?xml version="1.0" encoding="UTF-8"?>

<!--Annex II of WIPO Standard ST.26, Document Type Definition (DTD) for Sequence Listing

This entity may be identified by the PUBLIC identifier:

PUBLIC "-//WIPO//DTD SEQUENCE LISTING 1.3//EN" "ST26SequenceListing_V1_3.dtd"

* PUBLIC DTD URL

* https://www.wipo.int/standards/dtd/ST26SequenceListing_V1_3.dtd

Revision of Annex II to WIPO Standard ST.26 was approved by the Committee on WIPO Standards (CWS) at its sixth session.

* CONTACTS

xml.standards@wipo.int

* NOTES

The sequence data part is a subset of the complete INSDC DTD V.1.5 that only covers the requirements of WIPO Standard ST.26.

* REVISION HISTORY

2020-05-20: Version 1.3 approved at CWS/8.

Changes:

- Optional originalFreeTextLanguageCode attribute added to to allow applicants to indicate the language of the free text in the original sequence listing.
- Optional nonEnglishFreeTextLanguageCode attribute added to to allow applicants to indicate the language of the free text provided in the element .
- Optional id attribute added to INSDQualifier to facilitate comparison of language - dependent qualifier values between sequence listings.
- Optional element added to element to allow applicants to type language-dependent qualifiers in a non-English Language with the characters set forth in paragraph 40(a) of the ST.26 main body document.

2018-10-19: Version 1.2 approved at CWS/6.

Changes:

<INSDQualifier*> changed to <INSDQualifier+> for alignment with business needs and advice from NCBI (an INSDFeature_qual element (if present) should have one or more INSDQualifier elements).

2017-06-02: Version 1.1 approved at the CWS/5

Changes:

Comments added to <INSDSeq_length>, <INSDSeq_division> and <INSDSeq_sequence> to clarify the reason of the differences between the INSDC DTD v.1.5 and ST26 Sequence Listing DTD V1_1.

2016-03-24: Version 1.0 adopted at the CWS/4Bis

2014-03-11: Final draft for adoption.

ST26SequenceListing

* ROOT ELEMENT

-->

```
<!ELEMENT ST26SequenceListing ((ApplicantFileReference | (ApplicationIdentification,
ApplicantFileReference?)), EarliestPriorityApplicationIdentification?, (ApplicantName,
ApplicantNameLatin?)?, (InventorName, InventorNameLatin?)?, InventionTitle+, SequenceTotalQuantity,
SequenceData+)>
```

<!--The elements ApplicantName and InventorName are optional in this DTD to facilitate the conversion between various encoding schemes-->

<!--originalFreeTextLanguageCode:

The language code (see reference in paragraph 9 to ISO 639-1:2002) for the single original language in which the language-dependent free text qualifiers (NonEnglishQualifier_value) were prepared.

-->

<!--nonEnglishFreeTextLanguageCode:

The language code (see reference in paragraph 9 to ISO 639-1:2002) for the language in which the language-dependent free text qualifiers (NonEnglishQualifier_value) currently correspond.

-->

<!ATTLIST ST26SequenceListing

 dtdVersion CDATA #REQUIRED

 fileName CDATA #IMPLIED

 softwareName CDATA #IMPLIED

 softwareVersion CDATA #IMPLIED

 productionDate CDATA #IMPLIED

 originalFreeTextLanguageCode CDATA #IMPLIED

 nonEnglishFreeTextLanguageCode CDATA #IMPLIED

>

<!--ApplicantFileReference

Applicant's or agent's file reference, mandatory if application identification not provided.

-->

```
<!ELEMENT ApplicantFileReference (#PCDATA)>
```

<!--ApplicationIdentification

Application identification for which the sequence listing is submitted, when available.

-->

```
<!ELEMENT ApplicationIdentification (IPOfficeCode, ApplicationNumberText, FilingDate?)>
```

<!--EarliestPriorityApplicationIdentification

Identification of the earliest priority application, which contains IPOfficeCode, ApplicationNumberText and FilingDate elements.

-->

```
<!ELEMENT EarliestPriorityApplicationIdentification (IPOfficeCode, ApplicationNumberText,
FilingDate?)>
```

<!--ApplicantName

The name of the first mentioned applicant in characters set forth in paragraph 40(a) of the ST.26 main body document.

-->

<!--languageCode: Appropriate language code from ISO 639-1-Codes for the representation of names of languages - Part 1: Alpha-2

-->

```
<!ELEMENT ApplicantName (#PCDATA)>
```

<!ATTLIST ApplicantName

 languageCode CDATA #REQUIRED

>

<!--ApplicantNameLatin

Where ApplicantName is typed in characters other than those as set forth in paragraph 40(b), a translation or transliteration of the name of the first mentioned applicant must also be typed in characters as set forth in paragraph 40(b) of the ST.26 main body document.

```
-->
<!ELEMENT ApplicantNameLatin (#PCDATA)>
<!--InventorName
Name of the first mentioned inventor typed in the characters as set forth in paragraph 40(a).-->
<!--languageCode: Appropriate language code from ISO 639-1-Codes for the representation of names of
languages - Part 1: Alpha-2
-->
<!ELEMENT InventorName (#PCDATA)>
<!ATTLIST InventorName
        languageCode CDATA #REQUIRED
>
<!--InventorNameLatin
Where InventorName is typed in characters other than those as set forth in paragraph 40(b), a
translation or transliteration of the first mentioned inventor may also be typed in characters as set
forth in paragraph 40(b).
-->
<!ELEMENT InventorNameLatin (#PCDATA)>
<!--InventionTitle
Title of the invention typed in the characters as set forth in paragraph 40(a) in the language of
filing. A translation of the title of the invention into additional languages may be typed in the
characters as set forth in paragraph 40(a) using additional InventionTitle elements. The title of
invention is should be between two to seven words.
-->
<!--languageCode: Appropriate language code from ISO 639-1 - Codes
for the representation of names of languages - Part 1: Alpha-2
-->
<!ELEMENT InventionTitle (#PCDATA)>
<!ATTLIST InventionTitle
        languageCode CDATA #REQUIRED
>
<!--SequenceTotalQuantity
Indicates the total number of sequences in the document.
Its purpose is to be quickly accessible for automatic processing.
-->
<!ELEMENT SequenceTotalQuantity (#PCDATA)>
<!--SequenceData
Data for individual Sequence.
For intentionally skipped sequences see the ST.26 main body document.
-->
<!ELEMENT SequenceData (INSDSeq)>
<!ATTLIST SequenceData
        sequenceIDNumber CDATA #REQUIRED
>
<!--IPOfficeCode
ST.3 code. For example, if the application identification is PCT/IB2013/099999, then IPOfficeCode
value will be "IB" for the International Bureau of WIPO.
-->
<!ELEMENT IPOfficeCode (#PCDATA)>
<!--ApplicationNumberText
The application identification as provided by the office of filing (e.g. PCT/IB2013/099999)
-->
<!ELEMENT ApplicationNumberText (#PCDATA)>
<!--FilingDate
The date of filing of the patent application for which the sequence listing is submitted in ST.2
format "CCYY-MM-DD", using a 4-digit calendar year, a 2-digit calendar month and a 2-digit day within
the calendar month, e.g., 2015-01-31. For details, please see paragraphs 7 (a) and 11 of WIPO Standard
ST.2.
```

-->

```

<!ELEMENT FilingDate (#PCDATA)>
<!--*****
* INSD Part
*****

```

The purpose of the INSD part of this DTD is to define a customized DTD for sequence listings to support the work of IP offices while facilitating the data exchange with the public repositories.

The INSD part is subset of the INSD DTD v1.5 and as such can only be used to generate an XML instance as it will not support the complete INSD structure.

This part is based on:

The International Nucleotide Sequence Database (INSD) collaboration.

INSDSeq provides the elements of a sequence as presented in the GenBank/EMBL/DDBJ-style flatfile formats. Not all elements are used here.

-->

```

<!--INSDSeq
Sequence data. Changed INSD V1.5 DTD elements, INSDSeq_division and INSDSeq_sequence from optional to mandatory per business requirements.

```

-->

```

<!ELEMENT INSDSeq (INSDSeq_length, INSDSeq_moltype, INSDSeq_division, INSDSeq_other-seqids?, INSDSeq_feature-table?, INSDSeq_sequence)>

```

```

<!--INSDSeq_length

```

The length of the sequence. INSDSeq_length allows only integer.

-->

```

<!ELEMENT INSDSeq_length (#PCDATA)>

```

```

<!--INSDSeq_moltype

```

Admissible values: DNA, RNA, AA

-->

```

<!ELEMENT INSDSeq_moltype (#PCDATA)>

```

```

<!--INSDSeq_division

```

Indication that a sequence is related to a patent application. Must be populated with the value PAT.

-->

```

<!ELEMENT INSDSeq_division (#PCDATA)>

```

```

<!--INSDSeq_other-seqids

```

In the context of data exchange with database providers, the IPOs should populate for each sequence the element INSDSeq_other-seqids with one INSDSeqid containing a reference to the corresponding published patent and the sequence identification.

-->

```

<!ELEMENT INSDSeq_other-seqids (INSDSeqid?)>

```

```

<!--INSDSeq_feature-table

```

Information on the location and roles of various regions within a particular sequence. Whenever the element INSDSeq_feature-table is used, it must contain at least one feature.

-->

```

<!ELEMENT INSDSeq_feature-table (INSDFeature+)>

```

```

<!--INSDSeq_sequence

```

The residues of the sequence. The sequence must not contain numbers, punctuation or whitespace characters.

-->

```

<!ELEMENT INSDSeq_sequence (#PCDATA)>

```

```

<!--INSDSeqid

```

Intended for the use of IPOs in data exchange only.

Format:

```

pat|{office code}|{publication number}|{document kind code}|{Sequence identification number}

```

where office code is the code of the IP office publishing the patent document, publication number is the publication number of the application or patent, document kind code is the letter codes to distinguish patent documents as defined in ST.16 and Sequence identification number is the number of the sequence in that application or patent

Example:

```
pat|WO|2013999999|A1|123456
```

This represents the 123456th sequence from WO patent publication No. 2013999999 (A1)

```
-->
<!ELEMENT INSDSeqid (#PCDATA)>
<!--INSDFeature
Description of one feature.
-->
<!ELEMENT INSDFeature (INSDFeature_key, INSDFeature_location, INSDFeature_qual?)>
<!--INSDFeature_key
A word or abbreviation indicating a feature.
-->
<!ELEMENT INSDFeature_key (#PCDATA)>
<!--INSDFeature_location
Region of the presented sequence which corresponds to the feature.
-->
<!ELEMENT INSDFeature_location (#PCDATA)>
<!--INSDFeature_qual
List of qualifiers containing auxiliary information about a feature.
-->
<!ELEMENT INSDFeature_qual (INSDQualifier+)>
<!--INSDQualifier
Additional information about a feature.
For coding sequences and variants see the ST.26 main body document.
-->
<!--id
Unique identifier for the INSDQualifier to facilitate comparison of versions of a sequence listing
specifically having language-dependent qualifier values in different languages.
-->
<!ELEMENT INSDQualifier (INSDQualifier_name, INSDQualifier_value?,
NonEnglishQualifier_value?)>
<!ATTLIST INSDQualifier
            id ID #IMPLIED
>
<!--INSDQualifier_name
Name of the qualifier.
-->
<!ELEMENT INSDQualifier_name (#PCDATA)>
<!--INSDQualifier_value
Value of the qualifier. Where the qualifier is language dependent its value must be in the English
language and typed with the characters set forth in paragraph 40 (b).
-->
<!ELEMENT INSDQualifier_value (#PCDATA)>
<!--NonEnglishQualifier_value
Value of a language-dependent qualifier in a language that is not English and typed with the
characters set forth in paragraph 40 (a). The language is indicated with the attribute
nonEnglishFreeTextLanguageCode.
-->
<!ELEMENT NonEnglishQualifier_value (#PCDATA)>
```

[次頁から附属書Ⅲが続く]

附属書Ⅲ

配列表の実例(XML ファイル)

Version 1.4

2021 年 11 月 5 日開催の第 9 回 WIPO 標準委員会(CWS) 会合で採択

附属書Ⅲは以下のサイトで閲覧可能。：

https://www.wipo.int/standards/en/xml_material/st26/st26-annex-iii-sequence-listing-specimen.xml

[次頁から附属書Ⅳが続く]

附属書IV

配列表の XML インスタンスに用いる

Unicode 基本ラテンコード表の文字サブセット

Version 1.3

2020 年 12 月 4 日開催の第 8 回 WIPO 標準委員会(CWS) 会合で採択

アンパサンド文字(&のこと)(0026)は、予め定義された文字実体の一部としてのみ許可される。引用符(0022)、アポストロフィ(0027)、不等号(小号、003C)および不等号(大号、003E)は、それぞれ定義済み実体で表示しなければならない。また、アンパサンド文字(&のこと)(0026)は、属性値や要素の内容の中でアンパサンドとして使用される場合には、予め定義された文字実体で表されなければならない。

| Unicode code point | Character | Name |
|--------------------|-----------|------------------------|
| 0020 | | SPACE |
| 0021 | ! | EXCLAMATION MARK |
| 0022 | " | QUOTATION MARK |
| 0023 | # | NUMBER SIGN |
| 0024 | \$ | DOLLAR SIGN |
| 0025 | % | PERCENT SIGN |
| 0026 | & | AMPERSAND |
| 0027 | ' | APOSTROPHE |
| 0028 | (| LEFT PARENTHESIS |
| 0029 |) | RIGHT PARENTHESIS |
| 002A | * | ASTERISK |
| 002B | + | PLUS SIGN |
| 002C | , | COMMA |
| 002D | - | HYPHEN-MINUS |
| 002E | . | FULL STOP |
| 002F | / | SOLIDUS |
| 0030 | 0 | DIGIT ZERO |
| 0031 | 1 | DIGIT ONE |
| 0032 | 2 | DIGIT TWO |
| 0033 | 3 | DIGIT THREE |
| 0034 | 4 | DIGIT FOUR |
| 0035 | 5 | DIGIT FIVE |
| 0036 | 6 | DIGIT SIX |
| 0037 | 7 | DIGIT SEVEN |
| 0038 | 8 | DIGIT EIGHT |
| 0039 | 9 | DIGIT NINE |
| 003A | : | COLON |
| 003B | ; | SEMICOLON |
| 003C | < | LESS-THAN-SIGN |
| 003D | = | EQUALS SIGN |
| 003E | > | GREATER-THAN-SIGN |
| 003F | ? | QUESTION MARK |
| 0040 | @ | COMMERCIAL AT |
| 0041 | A | LATIN CAPITAL LETTER A |
| 0042 | B | LATIN CAPITAL LETTER B |
| 0043 | C | LATIN CAPITAL LETTER C |
| 0044 | D | LATIN CAPITAL LETTER D |
| 0045 | E | LATIN CAPITAL LETTER E |
| 0046 | F | LATIN CAPITAL LETTER F |
| 0047 | G | LATIN CAPITAL LETTER G |
| 0048 | H | LATIN CAPITAL LETTER H |
| 0049 | I | LATIN CAPITAL LETTER I |
| 004A | J | LATIN CAPITAL LETTER J |

| Unicode code point | Character | Name |
|--------------------|-----------|------------------------|
| 004B | K | LATIN CAPITAL LETTER K |
| 004C | L | LATIN CAPITAL LETTER L |
| 004D | M | LATIN CAPITAL LETTER M |
| 004E | N | LATIN CAPITAL LETTER N |
| 004F | O | LATIN CAPITAL LETTER O |
| 0050 | P | LATIN CAPITAL LETTER P |
| 0051 | Q | LATIN CAPITAL LETTER Q |
| 0052 | R | LATIN CAPITAL LETTER R |
| 0053 | S | LATIN CAPITAL LETTER S |
| 0054 | T | LATIN CAPITAL LETTER T |
| 0055 | U | LATIN CAPITAL LETTER U |
| 0056 | V | LATIN CAPITAL LETTER V |
| 0057 | W | LATIN CAPITAL LETTER W |
| 0058 | X | LATIN CAPITAL LETTER X |
| 0059 | Y | LATIN CAPITAL LETTER Y |
| 005A | Z | LATIN CAPITAL LETTER Z |
| 005B | [| LEFT SQUARE BRACKET |
| 005C | \ | REVERSE SOLIDUS |
| 005D |] | RIGHT SQUARE BRACKET |
| 005E | ^ | CIRCUMFLEX ACCENT |
| 005F | ˘ | LOW LINE |
| 0060 | ˘ | GRAVE ACCENT |
| 0061 | a | LATIN SMALL LETTER A |
| 0062 | b | LATIN SMALL LETTER B |
| 0063 | c | LATIN SMALL LETTER C |
| 0064 | d | LATIN SMALL LETTER D |
| 0065 | e | LATIN SMALL LETTER E |
| 0066 | f | LATIN SMALL LETTER F |
| 0067 | g | LATIN SMALL LETTER G |
| 0068 | h | LATIN SMALL LETTER H |
| 0069 | i | LATIN SMALL LETTER I |
| 006A | j | LATIN SMALL LETTER J |
| 006B | k | LATIN SMALL LETTER K |
| 006C | l | LATIN SMALL LETTER L |
| 006D | m | LATIN SMALL LETTER M |
| 006E | n | LATIN SMALL LETTER N |
| 006F | o | LATIN SMALL LETTER O |
| 0070 | p | LATIN SMALL LETTER P |
| 0071 | q | LATIN SMALL LETTER Q |
| 0072 | r | LATIN SMALL LETTER R |
| 0073 | s | LATIN SMALL LETTER S |
| 0074 | t | LATIN SMALL LETTER T |
| 0075 | u | LATIN SMALL LETTER U |
| 0076 | v | LATIN SMALL LETTER V |
| 0077 | w | LATIN SMALL LETTER W |
| 0078 | x | LATIN SMALL LETTER X |
| 0079 | y | LATIN SMALL LETTER Y |
| 007A | z | LATIN SMALL LETTER Z |
| 007B | { | LEFT CURLY BRACKET |
| 007C | | VERTICAL LINE |
| 007D | } | RIGHT CURLY BRACKET |
| 007E | ~ | TILDE |

[次頁から附属書 V が続く]

附属書 V

データ交換の追加要件（知的財産庁(IPOs)のみ参照）

Version 1.4

2021 年 11 月 5 日開催の第 9 回 WIPO 標準委員会(CWS) 会合で採択

データベースプロバイダ（国際塩基配列データベース(INSD)の会員）とのデータ交換において、知的財産庁は、配列ごとに、対応する公開特許への言及と配列番号を含む 1 つの要素 INSDSeqid を、次の形式で要素 INSDSeq_other-seqids に入力すべきである。

pat|{office code}|{publication number}|{document kind code}|{sequence identification number}

ここで、オフィスコードとは、ST.3 に規定されている特許文献を公開した産業財産権を所管する機関のコードであり、文書種類コードは、ST. 16 に規定されている異なる種類の特許文献を識別するためのコードであり、公開番号とは、その出願または特許の公開番号であり、配列番号とは、その出願または特許の中の配列の番号である。

例：

pat|WO|2013999999|A1|123456

これを有効な XML インスタンスに変換すると次のようになる。

```
<INSDSeq_other-seqids>  
  <INSDSeqid>pat|WO|2013999999|A1|123456</INSDSeqid>  
</INSDSeq_other-seqids>
```

ここで「123456」は、WIPO 発行の公開番号 2013999999 (A1)の 123456 番目の配列である。

[次頁から附属書 VI が続く]

附属書VI

実例付きガイダンス文書

Version 1.5

2021年11月5日開催の第9回WIPO標準委員会(CWS)会合で採択

目次

| | |
|-------|------------|
| はじめに | 3.26.vi.2 |
| 実例の索引 | 3.26.vi.7 |
| 実例 | 3.26.vi.17 |
| 付録 | 3.26.vi.74 |

はじめに

本標準は、その目的の一つとして、「出願人が、国際的および国内または広域的な手続きを目的として、特許出願で許容される単一の配列表を作成できるようにする」ことを示している。本ガイダンス文書の目的は、この目的を実現するために、すべての出願人および知的財産庁(IPO)が、配列を開示する際の記載と表示に関する要件を理解し、合意することである。

このガイダンス文書は、「はじめに」、「実例の索引」、「配列の開示例」および、「実例の配列をXML形式で記載した配列表を含む付録」で構成されている。この「はじめに」では、本ガイダンス文書で以下使用されている特定の概念および用語について説明する。事例では、本標準の特定の項の要件を説明しており、各事例には最も関連性のある項番号を表示している。一部の事例では他の項を詳細に説明しており、各事例の最後に他所参照を表示している。事例の索引には、各事例の頁番号と、所定の他所参照を表示している。配列表に記載しなければならない、あるいは任意で記載しても良い事例の各配列には、配列番号(SEQ ID NO)が割り当てられ、XML形式で本文書の付録に掲載している。

各事例に関し、配列と共に提示された記述的情報は、その配列に関する開示全体とみなされることを意図している。これは、事例で明示的に示された情報のみが考慮されることを意味している。

このガイダンス文書の説明は、特許出願日に提供するための配列表の作成を対象としており、提供された情報がIPOにより前の出願に主題事項を追加するとみなされる可能性がある関係上、特許出願日後に提供される配列表には適用できない場合がある。

配列表の作成

特許出願のための配列表作成について、以下の点を考慮する必要がある。

1. 本標準(ST. 26) 第7項は、特定の開示された配列の記載を求めているか？
2. 特定の開示された配列の記載が求められていない場合、その配列を記載することは ST.26 で認められているか？
3. 特定の開示された配列を記載することが ST.26 で求められている、あるいは認められている場合、その配列をどの様に配列表に記載すべきか？

最初の質問に関しては、ST. 26 第 7 項で、一定の制限付きで、特許出願で開示された配列が 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、その残基を数え上げて記載することが求められている。

2 番目の質問に関しては、ST. 26 第 8 項で、10 個未満の具体的に定義されたヌクレオチドまたは、4 個未満の具体的に定義されたアミノ酸からなる配列の配列表への記載を禁止している。

これら 2 つの質問に答えるためには、「残基の数え上げ」および「具体的に定義された」を明確に理解する必要がある。

3 番目の質問に関して、本文書では、様々なシナリオを例示した配列の開示と、各配列の好ましい表示方法、または配列に複数の変異がある場合は、「最も包括的な配列」に関する詳細な説明を示す。あらゆる配列のシナリオに対応することは不可能であるため、本ガイダンス文書では例示のない他の配列のシナリオにも同様の理由を適用することができるように、各例示の論拠と ST.26 の規定の適用方法を示すようにしている。

残基の数え上げ

ST. 26 第 3 項(c)では「残基の数え上げ」を、特許出願において配列の各残基を順に数え上げて開示することと定義しているが、その際、(i) 残基は、名称、略語、記号または構造により示されるか、(ii) 複数の残基が略式で示される。配列は、ST. 26 附属書 I、第 1 節、表 1 に記載されているヌクレオチド記号（すなわち小文字の記号または、大文字の同等物¹）および、ST.26 附属書 I、第 3 節、表 3 に記載されているアミノ酸記号（すなわち大文字の記号または小文字の同等物¹）である準拠型の記号を用いて、「その残基の数え上げ」によって特許出願に表示されるべきである。以下、これらのヌクレオチドおよびアミノ酸の記号を準拠型の記号と称し、表に記載されていないヌクレオチドおよびアミノ酸の表現を「非準拠型」と称する。

残基の表現が、準拠型の記号や略語（例えば、「Z1」は「A」を意味する）、あるいは準拠型の特定の配列（例えば、「Z1」は「agga」を意味する）に相当すると開示されている場合、ST.26 第 7 項が配列表に記載することを求めているか、あるいは第 8 項が記載することを禁止しているかを判断するために、その配列は、相当する準拠型の記号または略語を用いて表示されているかのように解釈される。非準拠型のヌクレオチド記号が曖昧さ記号として使用されているが（例えば、X1 = イノシンまたはプソイドウリジン）、第 1 節の表 1 に記載されている準拠型の曖昧さ記号のいずれか（すなわち、「m」、「r」、「w」、「s」、「y」、「k」、「v」、「h」、「d」、「b」、または「n」）に相当するものでない場合、その残基は「n」残基と解釈され、ST.26 第 7 項が配列表にその配列を記載することを求めているか、または ST.26 第 8 項が記載することを禁止するかを決定する。同様に、非準拠型のアミノ酸記号が曖昧さ記号として使用されているが（例えば、「Z1」は「A」、「G」、「S」または「T」を意味する）、第 3 節の表 3 の準拠型の曖昧さ記号（すなわち、B、Z、J、または X）のいずれかに相当するものでない場合、ST.26 第 7 項が配列表に配列を記載することを求めているか、または ST.26 第 8 項が記載することを禁止しているかを判断するために、その残基は「X」残基と解釈される。

準拠型の記号を使用しておよび配列を表示するように注意すべきであるが、配列が他の方法で表示されている場合、非準拠型の表現の意味を判断するために、説明の表示を考慮する必要がある場合がある。

準拠型の記号が使用されている場合には、その記号が準拠型の方法で使用されていることを確認するために、表示中の配列の説明を考慮しなければならない。記号が非準拠型の方法で使用されている場合は、ST. 26 第 7 項が配列表に記載することを求めているか、第 8 項が記載することを禁止しているかを判断するために、この説明をする必要がある。

具体的に定義された

ST.26 第 3 項(k)では、「具体的に定義された」とは、附属書 I に記載されている表中の記号「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であり、「n」および「X」は、第 1 節の表 1 に記載された（すなわち、「a」または「c」または「g」または「t/u」(「unknown」または「other」)) および第 3 節の表 3（すなわち、「A」、「R」、「N」、「D」、「C」、「Q」、「E」、「G」、「H」、「I」、「L」、「K」、「M」、「F」、「P」、「O」、「S」、「U」、「T」、「W」、「Y」、「V」のいずれか(「unknown」または「other」))に記載されている準拠型の方法で使用される。準拠型の記号または非準拠型の記号または略語、およびそれらの準拠型または非準拠型の方法での使用に関する上記の検討は、ヌクレオチドまたはアミノ酸が「具体的に定義されている」かどうかを判断するために考慮される。

最も包括的な配列

第 7 項の要件を満たす配列が、その残基を数え上げることで出願中に一度だけ表示されているが、複数の実施形態では異なる記載がなされている場合（例えば、ある実施形態では 1 個または複数の「X」は任意のアミノ酸である可能性があるが、別の実施形態では「X」は限られた数のアミノ酸のみである可能性がある場合、ST. 26 は、その残基によって数え上げられた単一配列のみを配列表に記載することを求めている。第 15 項および第 27 項のとおり、このような配列に複数の「n」または「x」の曖昧さ記号が含まれている場合、さらに注釈がない場合、「n」または「X」はそれぞれ任意のヌクレオチドまたはアミノ酸を表すと解釈される。その結果、配列表への記載を要する単一配列は、表示される配列の中で最も包括的な配列である。最も包括的な配列とは、最

¹ 注：本願開示の際、ヌクレオチドまたはアミノ酸を小文字または大文字の記号で表現することができるが、配列表に含まれる配列については、塩基配列の表現には小文字のみを使用しなければならない（ST.26 第 13 項参照）、アミノ酸配列の表現には大文字のみを使用しなければならない（ST.26 第 26 項参照）。

も実施形態の表示を含む最も限定的な曖昧さ記号によって表される変異した残基を有する単一配列である。同様に、配列がその残基の数え上げによって一度だけ表示されるが、配列の長さがコピー数多型によって変化する可能性がある場合、その配列の最も長い実施形態が最も包括的な配列であると考えられる。例えば、数え上げたように 2 から 5 コピーまで変化する反復領域を含む配列を考える。反復のコピー数が 5 個の実施形態が最も包括的な配列であり、配列表に記載すべきである。しかし、出来れば、本発明の主要な部分である追加の実施形態を表すなど、追加の特定の配列を記載することが強く推奨される。そうすることで、より綿密な検索が可能となり、特許を求める主題事項を公示することができる。

曖昧さ記号の利用

配列表内における曖昧さ記号「n」の適切な利用

記号「n」は、

- a. 単一のヌクレオチド以外のものを表すために使用してはならない。
- b. 説明情報と共に使用される場合を除き、「a」、「c」「g」または「t/u」のいずれかであると解釈する。
- c. 追加説明とともに、以下のヌクレオチドのいずれかを表すために使用するべきである。
 - i. 天然、合成、または非天然由来の修飾塩基で、附属書 I の他の記号では表現できないもの（第 1 節、表 1 参照）。
 - ii. 不明なヌクレオチド、すなわち、規定されていない、開示されていない、または不確かなヌクレオチド
 - iii. 脱塩基部位、または
- d. 配列変化、すなわち、代替、欠失、挿入、または置換を表すために使用しても良く、「n」は最も限定的な曖昧さ記号である。

記号「X」は、

- a. 単一のアミノ酸以外のものを表すために使用してはならない。
- b. 説明情報と共に使用される場合を除き、「A」、「R」、「N」、「D」、「C」、「Q」、「E」、「G」、「H」、「I」、「L」、「K」、「M」、「F」、「P」、「O」、「S」、「U」、「T」、「W」、「Y」または「V」のいずれかであると解釈する。
- c. 追加説明とともに、以下のアミノ酸のいずれかを表すために使用するべきである。
 - i. 天然、合成、または非天然由来の修飾アミノ酸で、附属書 I の他の記号では表現できないもの（第 3 節、表 3 参照）。
 - ii. 不明なアミノ酸、すなわち、規定されていない、開示されていない、または不確かなアミノ酸、または
- d. 配列変化、すなわち、代替、欠失、挿入、または置換を表すために使用しても良く、「X」は最も限定的な曖昧さ記号である。

修飾残基の注釈

本標準では、ヌクレオチドについては第 17 項、アミノ酸については第 30 項に基づいて、「修飾」残基に注釈を付すことが求められている。

ST.26 第 3 項 (e) では、第 3 項(a)の記載のアミノ酸のうち、「修飾アミノ酸」を、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-リジン、L-メチオニン、L-フェニルアラニン、L-プロリン、L-ピロリジン、L-セリン、L-セレンシステイン、L-トレオニン、L-トリプトファン、L-チロシン、または L-バリン以外のものと定義している。同様に、「修飾塩基」を、第 3 項 (g) に記載のヌクレオチドのうち、デオキシアデノシン 3'-ウリジン酸、デオキシグアノシン 3'-ウリジン酸、デオキシチミジン 3'-ウリジン酸、デオキシシチジン 3'-ウリジン酸、アデノシン 3'-ウリジン酸、グアノシン 3'-ウリジン酸、シチジン 3'-ウリジン酸、またはウリジン 3'-ウリジン酸以外のものと定義している (ST.26, 第 3 項(f))。

上記の定義に基づけば、核酸の核酸塩基または糖リン酸骨格の修飾、およびペプチドのアンモニウム酸 R 基またはペプチド骨格の修飾は、それぞれ 1 つまたは複数の「修飾塩基」または「修飾アミノ酸」をもたらす。したがって、そのようなヌクレオチドやアミノ酸には注釈を付す必要がある。骨格の修飾の例としては、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、および D-アミノ酸が挙げられる。

ペプチドの末端アミノ酸や核酸の末端ヌクレオチドを修飾しても、必ずしも「修飾アミノ酸」や「修飾塩基」になるわけではないことに注意を要する。末端の修飾を見て、その修飾が残基の化学構造を変え、残基が第 3 項(e)および第 3 項(f)の例外に該当する

かどうかを判断する必要がある。例えば、C 末端残基が準抛型のアミド結合を介して構造（分岐配列の一部など- 事例 7 (b) -3 のペプチド#2 を参照）に連結されているペプチドは、「修飾残基」とはみなされず、したがって注釈を付す必要はない。同様に、N 末端残基がビオチンにアミド結合しているペプチドは、「修飾残基」とはみなされず、したがって、注釈をつける必要はない。どちらのシナリオでも、C 末端または N 末端の結合に関わる残基の構造は、本標準の第 3 項(e)に記載されている準抛型のアミノ酸から変更されていない。

対照的に、残基の化学構造を変える末端修飾は「修飾残基」とみなされ、注釈を付さなければならない。例えば、実施例 3(c)-1 の C 末端のメチル化は、 α カルボキシル基に通見られる水酸基をメチル基が置換するため、末端残基の化学構造を変化させる。したがって、このメチル化されたリジンは「修飾残基」として注釈を付さなければならない。

なお、数え上げられた配列内の各末端残基の修飾を評価し、末端残基の構造が変化しているかどうかを判断するのは出願人に委ねられている。修飾残基の構造が、本標準の第 3 項(e)および第 3 項(f)に記載されている準抛型のアミノ酸またはヌクレオチドと異なる場合は、その修飾について注釈を付さなければならない。

最後に、出願人には、表示内容を可能な限り正確に表現するために、配列表に合理的な範囲で常に多くの情報を記載することを推奨する。したがって、注釈を付す必要がない修飾であっても、できれば記載すべきである。

ただし、数え上げられた一次配列の変異体の注釈は、ST.26 第 93 項から第 100 項の定義に準拠しなければならない。数え上げられた配列の変異体として表示された修飾は、配列表に記載する必要がない場合もある。変異体の注釈の定義については、ST.26 第 93 項から第 95 項 を参照のこと。

修飾残基の表現

ST.26 では、修飾塩基やアミノ酸は、可能な限り対応する非修飾残基として配列表に表記すべきであるとしている（第 16 項および第 29 項参照）。この提案事項は「すべき」であり、「強く推奨されるアプローチであるが、必須ではない」ことに留意されたい（第 4 項(d)参照）。修飾残基を、対応する非修飾残基で表すか、変数「n」または「X」で表すかは、出願人の裁量に委ねられている。

一般的な経験則として、ある残基がメチル化やアセチル化などの部位の付加によって修飾されていて、非修飾残基の構造が一般的に変化しない場合、非修飾残基で表現することが推奨される。例えば、メチル化されたアデノシンは、配列表では好ましくは「a」で表されるべきである。しかし、修飾残基が非修飾残基と構造的に異なる場合には、「n」または「X」を用いることが推奨される。例えば、ノルロイシンはロイシンの異性体であり、その側鎖は炭素数 4 の直鎖構造である。ロイシンも炭素数 4 の側鎖を有しているが、2 番目の炭素で分岐している。そのため、ノルロイシンは単にロイシンに修飾が加えられたものではなく、（関連はあるものの）全く異なる構造をしている。そのため、配列表ではノルロイシンを「X」で表すことが推奨される。

ヌクレオチドは「n」以外で表されると「具体的に定義される」ものであり、アミノ酸は「X」以外で表されると「具体的に定義される」ものである（ST.26、第 3 項(k)項参照）。したがって、配列中に「a」で表される 2'-O-メチルアデノシンは具体的に定義される一方、配列中の「X」で表されるノルロイシンは具体的に定義されない。

表 A – 準拠型ヌクレオチドの記号および定義

| Symbol | Definition |
|--------|---|
| 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” |

表 B- 準拠型アミノ酸の記号、3文字コードおよび定義

| Symbol | 3-Letter Code | Definition |
|--------|---------------|--|
| A | Ala | Alanine |
| R | Arg | Arginine |
| N | Asn | Asparagine |
| D | Asp | Aspartic Acid (Aspartate) |
| C | Cys | Cysteine |
| Q | Gln | Glutamine |
| E | Glu | Glutamic Acid (Glutamate) |
| G | Gly | Glycine |
| H | His | Histidine |
| I | Ile | Isoleucine |
| L | Leu | Leucine |
| K | Lys | Lysine |
| M | Met | Methionine |
| F | Phe | Phenylalanine |
| P | Pro | Proline |
| O | Pyl | Pyrrolysine |
| S | Ser | Serine |
| U | Sec | Selenocysteine |
| T | Thr | Threonine |
| W | Trp | Tryptophan |
| Y | Tyr | Tyrosine |
| V | Val | Valine |
| B | Asx | Aspartic Acid or Asparagine |
| Z | Glx | Glutamine or Glutamic Acid |
| J | Xle | Leucine or Isoleucine |
| X | Xaa | 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" |

事例の索引

第 3 項(a) – 「アミノ酸」の定義

例 3(a) -1: D-アミノ酸

他所参照 (クロスリファレンス)

例 29-1: 「その他」のアミノ酸を示す最も限定的な曖昧さ記号

例 29-2: 対応する非修飾アミノ酸の使用

例 30-1: feature key「CARBOHYD」

第 3 項(c) – 「残基の数え上げ」の定義

例 3(c) -1: 化学構造によるアミノ酸の数え上げ

例 3(c) -2: アミノ酸配列の簡略式

他所参照 (クロスリファレンス)

例 27-1: アミノ酸配列の簡略式

例 27-3: 簡略式– 4 つまたはそれ以上の具体的に定義されたアミノ酸

第 3 項(f) – 「修飾塩基」の定義他所参照 (クロスリファレンス)

例 3(g) -4: 核酸類似体

第 3 項(g) – 「ヌクレオチド」の定義

例 3(g) -1: C3 スペースで区切られた塩基配列

例 3(g) -2: C3 スペースを含む代替残基を有する塩基配列

例 3(g) -3: 脱塩基部位

例 3(g) -4: 核酸類似体

他所参照 (クロスリファレンス)

例 11(b) -1: 二本鎖の塩基配列 - 長さが異なる場合

例 14-1: RNA 中のウラシルを示す記号「t」

第 3 項(k) – 「具体的に定義された」の定義

例 3(k) -1: ヌクレオチドに用いる曖昧さ記号

例 3(k) -2: 準拠型の方法と非準拠型の方法の両方で使用される曖昧さ記号「n」

例 3(k) -3: 非準拠型の方法で使用される曖昧さ記号「n」

例 3(k) -4: 「n」以外の曖昧さ記号が「具体的に定義された」残基を示している場合

例 3 (k) -5: 非準拠型の方法で使用される曖昧さ記号「Xaa」

第 7 項 - 配列表への記載が求められている配列他所参照 (クロスリファレンス)

- 例 28-1: コードする塩基配列、およびコードされたアミノ酸配列
- 例 55-1: DNA/RNA 分子の結合
- 例 89-2: 開示されている配列を超える feature location
- 例 92-1: イントロンを有するコード配列によってコードされたアミノ酸配列

第 7 項(a) - 配列表に必要な塩基配列

- 例 7(a) -1: 分岐塩基配列
- 例 7(a) -2: 二次構造を有する線形の塩基配列
- 例 7(a) -3: 非準拠型の方法で使用されるヌクレオチドの曖昧さ記号
- 例 7(a) -4: 非準拠型の方法で使用されるヌクレオチドの曖昧さ記号
- 例 7(a) -5: 非準拠型ヌクレオチドの記号
- 例 7(a) -6: 非準拠型ヌクレオチドの記号

他所参照 (クロスリファレンス)

- 例 3(g) -1: C3 スペースで区切られた塩基配列
- 例 3(g) -2: C3 スペースを含む代替残基を有する塩基配列
- 例 3(g) -3: 脱塩基部位
- 例 3(g) -4: 核酸類似体
- 例 3(k) -1: ヌクレオチドに用いる曖昧さ記号
- 例 3(k) -2: 準拠型の方法と非準拠型の方法の両方で使用される曖昧さ記号「n」
- 例 3(k) -3: 非準拠型の方法で使用される曖昧さ記号「n」
- 例 3(k) -4: 「n」以外の曖昧さ記号が「具体的に定義された」場合
- 例 11(a) -1: 二本鎖の塩基配列 - 長さが同じもの
- 例 11(b) -1: 二本鎖の塩基配列 - 長さが異なるもの
- 例 11(b) -2: 二本鎖の塩基配列 - 塩基対のセグメントを持たないもの
- 例 14-1: RNA 中のウラシルを示す記号「t」
- 例 89-1: コードする塩基配列、およびコードされたアミノ酸配列
- 例 93-1: 数え上げられた変異体の表示
- 例 95(b) -1: 複数の相互に依存する変異体を有する、個々の変異体配列の表示

第 7 項(b) - 配列表に必要なアミノ酸配列

例 7(b) -1: 4 つまたはそれ以上の具体的に定義されたアミノ酸

例 7(b) -2: 分岐アミノ酸配列

例 7(b) -3: 分岐アミノ酸配列

例 7(b) -4: 分岐アミノ酸配列を含む環状ペプチド

例 7(b) -5: 分岐アミノ酸配列を含む環状ペプチド

他所参照 (クロスリファレンス)

例 3(a)-1: D-アミノ酸

例 3(c) -1: 化学構造によるアミノ酸の数え上げ

例 3(c) -2: アミノ酸配列の簡略式

例 3 (k) -5: 非準拠型の方法で使用される曖昧さ記号「Xaa」

例 27-1: 塩基配列の簡略式

例 27-3: 簡略式 – 4 つまたはそれ以上の具体的に定義されたアミノ酸

例 29-1: 「その他」のアミノ酸を示す最も限定的な曖昧さ記号

例 29-2: 対応する非修飾アミノ酸の使用

例 30-1: feature key「CARBOHYD」

例 36-1: 単一配列として示される既知数の「X」残基の領域を 1 つ有する配列

例 37-1: 単一配列として示してはならない、未知数の「X」残基の領域を有する配列

例 37-2: 単一配列として示してはならない、未知数の「X」残基の領域を有する配列

例 89-1: コードする塩基配列、およびコードされたアミノ酸配列

例 93 -2: 数え上げられた変異体の表示

例 93 -3: 共通配列の表示

例 94 -1: 数え上げられた代替アミノ酸による単一配列の表現

例 95(a) -1: 一次配列の注釈による変異体配列の表示

第 8 項 - 配列の記載に用いる閾値

他所参照 (クロスリファレンス)

例 3(k) -1: ヌクレオチドに用いる曖昧さ記号

例 3(k) -2: 準拠型の方法と非準拠型の方法の両方で使用される曖昧さ記号「n」

例 7(a) -1: 分岐塩基配列

例 7(a) -6: 非準拠型ヌクレオチドの記号

例 7(b) -1: 4 つまたはそれ以上の具体的に定義されたアミノ酸

例 7(b) -2: 分岐アミノ酸配列

例 7(b) -4: 分岐アミノ酸配列を含む環状ペプチド

例 14-1: RNA 中のウラシルを示す記号「t」

例 37-1: 単一配列として示してはならない、未知数の「X」残基の領域を有する配列

例 37-2: 単一配列として示してはならない、未知数の「X」残基の領域を有する配列

例 94 -1: 数え上げられた代替アミノ酸による単一配列の表示

第 11 項 - 塩基配列の表示

他所参照 (クロスリファレンス)

例 3(g) -4: 核酸類似体

例 7(a) -1: 分岐塩基配列

第 11 項(a) - 二本鎖の塩基配列 - 完全に補完しあう塩基配列

例 11(a) -1: 二本鎖の塩基配列 - 長さが同じもの

第 11 項(b) - 二本鎖の塩基配列 - 完全に補完しあう塩基配列ではない塩基配列

例 11(b) -1: 二本鎖の塩基配列 - 長さが異なるもの

例 11(b) -2: 二本鎖の塩基配列 - 塩基対のセグメントを持たないもの

第 13 項 - ヌクレオチドの表示

他所参照 (クロスリファレンス)

例 3(k) -2: 準拠型の方法、非準拠型の方法両方で使用される曖昧さ記号「n」

例 7(a) -1: 分岐塩基配列

例 14-1: RNA を構成する塩基のウラシルを示す記号「t」

例 93-1: 数え上げられた変異体の表示

第 14 項 - RNA 中のウラシルを示すと解釈される記号「t」

例 14-1: RNA 中のウラシルを示す記号「t」

他所参照 (クロスリファレンス)

例 55-1: DNA/RNA 分子の結合

第 15 項 - ヌクレオチドを示す最も限定的な曖昧さ記号を用いるべき場合

他所参照 (クロスリファレンス)

例 3(g) -1: C3 スペースで区切られた塩基配列

例 3(g) -2: C3 スペースを含む代替残基を有する塩基配列

例 3(k) -4: 「n」以外の曖昧さ記号が「具体的に定義された」残基を示している場合

例 95(b) -1: 複数の相互に依存するバリエーションを有する、個々の変異体配列の表示

第 16 項 - 修飾塩基の表示

他所参照 (クロスリファレンス)

例 3(g) -1: C3 スーパーで区切られた塩基配列

例 3(g) -4: 核酸類似体

第 17 項 - 修飾塩基の注釈

他所参照 (クロスリファレンス)

例 3(g) -1: C3 スーパーで区切られた塩基配列

例 3(g) -3: 脱塩基部位

例 7(a) -1: 分岐塩基配列

例 7(a) -2: 二次構造を有する線形の塩基配列

例 7(a) -6: 非準拠型ヌクレオチドの記号

第 18 項 - 連続した修飾塩基を有する領域の注釈

他所参照 (クロスリファレンス)

例 3(g) -4: 核酸類似体

例 11(b) -1: 二本鎖の塩基配列 - 長さが異なる場合

第 19 項 - DNA 中のウラシルもしくは RNA 中のチミンの注釈

他所参照 (クロスリファレンス)

例 14-1: RNA 中のウラシルを示す記号「t」

第 25 項 - 残基位置番号 1 のアミノ酸配列

他所参照 (クロスリファレンス)

例 3(a)-1: D-アミノ酸

例 7(b) -4: 分岐アミノ酸配列を含む環状ペプチド

例 7(b) -5: 分岐アミノ酸配列を含む環状ペプチド

例 29-1: 「その他」のアミノ酸を示す最も限定的な曖昧さ記号

第 26 項 - アミノ酸の表示

他所参照 (クロスリファレンス)

例 7(b) -2: 分岐アミノ酸配列

- 例 7(b) -4: 分岐アミノ酸配列を含む環状ペプチド
- 例 7(b) -5: 分岐アミノ酸配列を含む環状ペプチド
- 例 36-1: 単一配列として示される既知数の「X」残基の領域を 1 つ有する配列
- 例 89-1: コードする塩基配列、およびコードされたアミノ酸配列
- 例 92-1: イントロンを有するコード配列によってコードされたアミノ酸配列
- 例 93-2: 数え上げられた変異体の表示
- 例 93 -3: 共通配列の表示

第 27 項 - アミノ酸を示す最も限定的な曖昧さ記号を用いるべき場合

- 例 27-1: アミノ酸配列の簡略式
- 例 27-2: 簡略式 - 4 つ未満の具体的に定義されたアミノ酸
- 例 27-3: 簡略式 - 4 つ以上の具体的に定義されたアミノ酸

他所参照 (クロスリファレンス)

- 例 3(c) -2: アミノ酸配列の簡略式
- 例 7(b) -1: 4 つ以上の具体的に定義されたアミノ酸
- 例 29-1: 「その他」のアミノ酸を示す最も限定的な曖昧さ記号
- 例 36-1: 単一配列として示される既知数の「X」残基の領域を 1 つ有する配列
- 例 36-2: 単一配列として示される既知数の「X」残基の領域を複数有する配列
- 例 36-3: 単一配列として示される既知数の「X」残基の領域を複数有する配列
- 例 37-2: 単一配列として示してはならない、未知数の「X」残基の領域を有する配列
- 例 93 -3: 共通配列の表示
- 例 94 -1: 数え上げられた代替アミノ酸による単一配列の表現
- 例 95(a) -1: 一次配列の注釈による変異体配列の表示

第 28 項 - 内部終結部位記号によって区切られたアミノ酸配列

- 例 28-1: コードする塩基配列、およびコードされたアミノ酸配列
- ### 他所参照 (クロスリファレンス)
- 例 89-1: コードする塩基配列、およびコードされたアミノ酸配列
 - 例 92-1: イントロンを有するコード配列によってコードされたアミノ酸配列

第 29 項 - 「その他」のアミノ酸の表示

- 例 29-1: 「その他」のアミノ酸を示す最も限定的な曖昧さ記号
- 例 29-2: 対応する非修飾アミノ酸の使用

他所参照 (クロスリファレンス)

- 例 3(a)- 1: D-アミノ酸
- 例 7(b) -2: 分岐アミノ酸配列
- 例 7(b) -3: 分岐アミノ酸配列
- 例 7(b) -4: 分岐アミノ酸配列を含む環状ペプチド
- 例 30-1: feature key「CARBOHYD」

第 30 項 - 修飾アミノ酸の注釈**例 30-1: feature key「CARBOHYD」**他所参照 (クロスリファレンス)

- 例 3(a)- 1: D-アミノ酸
- 例 3(c) -1: 化学構造によるアミノ酸の数え上げ
- 例 7(b) -2: 分岐アミノ酸配列
- 例 7(b) -3: 分岐アミノ酸配列
- 例 7(b) -4: 分岐アミノ酸配列を含む環状ペプチド
- 例 7(b) -5: 分岐アミノ酸配列を含む環状ペプチド
- 例 29-1: 「その他」のアミノ酸を示す最も限定的な曖昧さ記号
- 例 29-2: 対応する非修飾アミノ酸の使用

第 31 項 - D-アミノ酸の表示他所参照 (クロスリファレンス)

- 例 3(a) -1: D-アミノ酸
- 例 3(c) -1: 化学構造によるアミノ酸の数え上げ
- 例 7(b) -2: 分岐アミノ酸配列
- 例 7(b) -3: 分岐アミノ酸配列
- 例 7(b) -4: 分岐アミノ酸配列を含む環状ペプチド
- 例 7(b) -5: 分岐アミノ酸配列を含む環状ペプチド

第 32 項 - 「不明な」アミノ酸の注釈他所参照 (クロスリファレンス)

- 例 3(c) -1: 化学構造によるアミノ酸の数え上げ

第 34 項 - 「X」残基の隣接する領域の注釈

他所参照 (クロスリファレンス)

例 29-1: 「その他」のアミノ酸を示す最も限定的な曖昧さ記号

第 36 項 - 正確な数の隣接する「n」または「X」残基の領域を含む配列

例 36-1: 単一配列として示される既知数の「X」残基の領域を 1 つ有する配列

例 36-2: 単一配列として示される既知数の「X」残基の領域を複数有する配列

例 36-3: 単一配列として示される既知数の「X」残基の領域を複数有する配列

第 37 項 - 隣接する「n」または「X」残基の数が不明な領域を含む配列

例 37-1: 単一配列として示してはならない、未知数の「X」残基の領域を有する配列

例 37-2: 単一配列として示してはならない、未知数の「X」残基の領域を有する配列

第 41 項 - 予約文字他所参照 (クロスリファレンス)

例 89-2: 開示された配列を超える feature location

第 54 項 - 要素 INSDSeq_moltype他所参照 (クロスリファレンス)

例 14-1: RNA 中のウラシルを示す記号「t」

第 55 項 - DNA と RNA の両方を含む塩基配列

例 55-1: DNA/RNA 分子の結合

第 56 項 - DNA セグメントと RNA セグメントの両方を含む塩基配列を示す例他所参照 (クロスリファレンス)

例 55-1: DNA/RNA 分子の結合

第 57 項 - 要素 INSDSeq_sequence他所参照 (クロスリファレンス)

例 28-1: コードする塩基配列、およびコードされたアミノ酸配列

例 92-1: イントロンを有するコード配列によってコードされたアミノ酸配列

第 65 項 - 存在位置記述子他所参照 (クロスリファレンス)

例 3(g) -4: 核酸類似体

例 89-2: 開示された配列を超える feature location

第 66 項 - 存在位置記述子の構文他所参照 (クロスリファレンス)

例 3(g) -4: 核酸類似体

例 7(b) -4: 分岐アミノ酸配列を含む環状ペプチド

例 29-1: 「その他」のアミノ酸を示す最も限定的な曖昧さ記号

例 30-1: feature key「CARBODHYD」

例 89-2: 開示された配列を超える feature location

第 67 項 – 存在位置演算子

他所参照 (クロスリファレンス)

例 92-1: イントロンを有するコード配列によってコードされたアミノ酸配列

第 70 項 – feature location

他所参照 (クロスリファレンス)

例 7(b)-4: 分岐アミノ酸配列を含む環状ペプチド

例 29-1: 「その他」のアミノ酸を示す最も限定的な曖昧さ記号

例 30-1: feature key「CARBODHYD」

例 89-2: 開示された配列を超える feature location

第 71 項 – 存在位置記述子における文字「<」および「>」の記載

他所参照 (クロスリファレンス)

例 29-1: 「その他」のアミノ酸を示す最も限定的な曖昧さ記号

例 89-2: 開示された配列を超える feature location

第 83 項 – 自然発生ではない塩基配列を示す例

他所参照 (クロスリファレンス)

例 55-1: DNA/RNA 分子の結合

第 89 項 – feature key「CDS」

例 89-1: コードする塩基配列、およびコードされたアミノ酸配列

例 89-2: 開示された配列を超える feature location

他所参照 (クロスリファレンス)

例 92-1: イントロンを有するコード配列によってコードされたアミノ酸配列

第 90 項 – qualifier 「transl table」 および 「translation」

他所参照 (クロスリファレンス)

例 28-1: コードする塩基配列、およびコードされたアミノ酸配列

例 89-1: コードする塩基配列、およびコードされたアミノ酸配列

例 92-1: イントロンを有するコード配列によってコードされたアミノ酸配列

第 92 項 - コード配列によりコードされたアミノ酸配列

例 92-1: イントロンを有するコード配列によってコードされたアミノ酸配列

他所参照 (クロスリファレンス)

例 28-1: コードする塩基配列、およびコードされたアミノ酸配列

例 89-1: コードする塩基配列、およびコードされたアミノ酸配列

例 89-2: 開示された配列を超える feature location

第 93 項 - それぞれ残基により数え上げられた、一次配列と変異体

例 93-1: 数え上げられた変異体の表示

例 93 -2: 数え上げられた変異体の表示

例 93 -3: 共通配列の表示

第 94 項 - 数え上げられた代替残基を有する単一配列として開示される変異体配列

例 94 -1: 数え上げられた代替アミノ酸による単一配列の表示

第 95 項(a) - 複数の独立した変異体を有する一次配列によってのみ開示される変異体配列

例 95(a) -1: 一次配列の注釈による変異体配列の表示

第 95 項(b) - 複数の相互に依存する変異体を有する一次配列によってのみ開示される変異体配列

例 95(b) -1: 複数の相互に依存する変異体を有する、個々の変異体配列の表示

第 96 項 - 変異体配列に関する feature key および qualifier

他所参照 (クロスリファレンス)

例 29-1: 「その他」のアミノ酸を示す最も限定的な曖昧さ記号

第 97 項 - 変異体配列の注釈

他所参照 (クロスリファレンス)

例 29-1: 「その他」のアミノ酸を示す最も限定的な曖昧さ記号

例 93 -3: 共通配列の表示

例 94 -1: 数え上げられた代替アミノ酸による単一配列の表示

事 例

第 3 項(a) 「アミノ酸」の定義

例 3(a) -1: D-アミノ酸

以下の配列が記載された特許出願：

Cyclo (D-Ala-D-Glu-Lys-Nle-Gly-D-Met-D-Nle)

質問 1：ST. 26 では上記配列の記載を要するか？

回答：はい

ST. 26 の第 3 項(a)で「アミノ酸」は、「D-アミノ酸」および修飾を受けたまたは合成の側鎖を含むアミノ酸が含まれると定義されている。この定義に基づいて、連続したペプチドは、具体的に定義された 5 つのアミノ酸 (D-Ala, D-Glu, Lys, Gly, and D-Met) を含んでいる。したがって、この配列は ST. 26 の第 7 項(b)で求められているように、配列表に記載しなければならない。

質問 3：この配列を配列表にどの様に記載すべきか？

第 29 項では、D-アミノ酸は、対応する非修飾の L-アミノ酸と同じ配列で表されるべきであるとしている。さらに、附属書 I の第 3 節、表 3 の他の記号で表すことができない修飾を受けたアミノ酸は、記号「X」で表さなければならない。

この例では、配列には、附属書 I、第 3 節、表 3 の非修飾の L-アミノ酸で表すことができる 3 つの D-アミノ酸、1 つの L-アミノ酸 (Nle)、および記号「X」で表さなければならない 1 つの D-アミノ酸 (D-Nle) が含まれている。

第 25 項は、アミノ酸配列が環状構造で、その環はペプチド結合で結合されたアミノ酸残基のみで構成されている場合、出願人は残基位置番号 1 のアミノ酸を選択しなければならないことを示している。したがって、この配列は次のように表示できる：

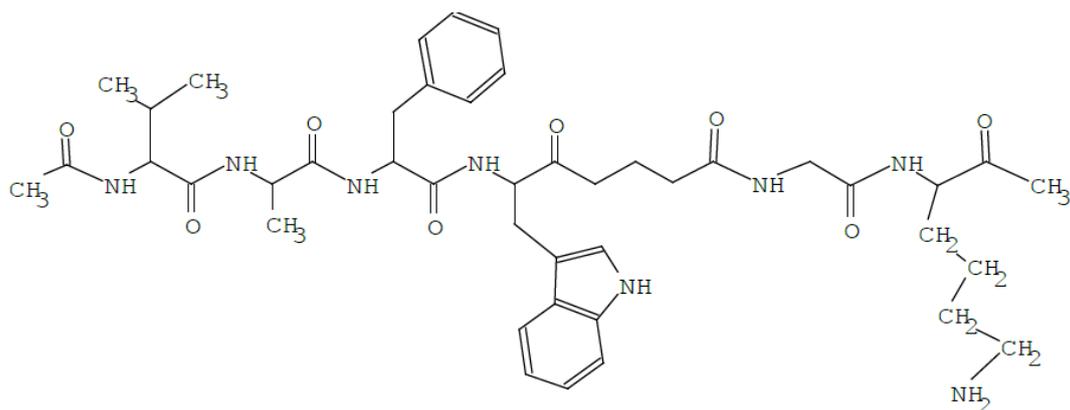
AEKXGMX (配列番号 1)

また、残基位置番号 1 に、配列内の他のアミノ酸と一緒に表示している場合もある。feature key「SITE」および qualifier「note」は、各 D-アミノ酸に対して、D-アミノ酸の省略されていない完全名前を qualifier の値として提供しなければならない。さらに、feature key「SITE」および qualifier「note」には、附属書 I のセクション 4、表 4 に記載されているように、qualifier として L-ノルロイシンの略語「Nle」を使用しなければならない。最後に、ペプチドが環状であることを示すために、feature key「REGION」および qualifier「note」を提供すべきである。

関連する ST. 26 の項：第 3 項(a)、第 7 項(b)、第 25 項、第 26 項、第 29 項、第 30 項および第 31 項

第 3 項(c) – 「残基の数え上げ」

例 3(c) -1: 化学構造によるアミノ酸の数え上げ



質問 1: ST. 26 では上記配列の記載を要するか?

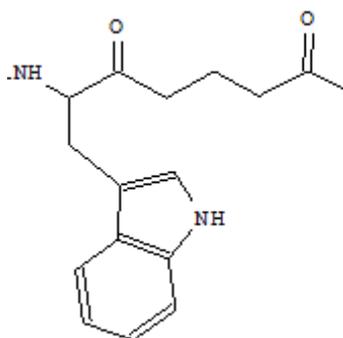
回答: はい

構造式として上で示されている連続したペプチドは、少なくとも 4 つの具体的に定義されたアミノ酸が含まれており、したがって、この配列は配列表の中に含めなければならない。

質問 3: この配列を配列表にどのように記載すべきか?

配列は以下の通り記載しても良い。

VAFXGK (配列番号 2)



ここで、「X」は「その他」の修飾を受けたアミノ酸を示しており、feature key「SITE」および qualifier「note」が必須となる。qualifier「note」は、数え上げられたペプチドの 4 番目の位置にある修飾を受けたトリプトファン、省略されていない完全名称、例えば「6-amino-7-(1H-indol-3-yl)-5-oxoheptanoic acid」を qualifier 値として記載する。および C 末端がメチル化されると、末端の-OH が-CH₃ に置換されるため、末端のリジンの化学構造が変化する。この構造変化により、配列内のリジンは「修飾を受けたアンモニウム酸」とみなされる。そのため、C 末端のメチル化を示す feature key「SITE」および qualifier「note」が必要となる。一方、パリンは、アセチル基の付加が通常のペプチド結合を伴うため、「修飾アミノ酸」とはみなされない。また、アシル化によってパリンの構造が変化することもない。したがって、N 末端のアセチル化を示すために、追加の feature key「SITE」および qualifier「note」の記載を要する。

あるいは、この配列は以下の通り記載しても良い。

VAFW (配列番号 3)

数え上げられたペプチドの 4 位のトリプトファンが修飾されていることを示すために、feature key「SITE」と qualifier「note」が必要であり、「C-terminus linked via a glutaraldehyde bridge to dipeptide GK」という値で示される。さらに、N 末端のアセチル化を示すために、存在位置 1 に追加で feature key「SITE」および qualifier「note」の記載が必要である。

関連する ST. 26 の項: 第 3 項(c)、第 7 項(b)、第 29 項、第 30 項および第 31 項

例 3(c) -2: アミノ酸配列の簡略式

$(G_4z)_n$

上記の式中の G はグリシン、z は任意のアミノ酸、変数「n」は任意の整数とする。

質問 1：ST. 26 では上記配列の記載を要するか？

回答：はい

上記の簡略式中の数値「n」は任意の整数であることが示されており、したがって、「n」の最も包括的な実施形態は不定元である。「n」が不定元であるため、この簡略式で示されているペプチドを確定した長さに展開することができず、したがって、展開しない式を考慮しなければならない。

展開しない式（「n」= 1）の数え上げられたペプチドは、それぞれが Gly である 4 つの具体的に定義されたアミノ酸と、記号「z」を示している。従来、「Z」は「グルタミンまたはグルタミン酸」を表す記号であったが、本実施例では「Z」を「任意のアミノ酸」と定義している。ST.26 では、具体的に定義されていないアミノ酸は「X」で表される。この分析によれば、数え上げられたペプチド、すなわち GGGGX は、数え上げられ、具体的に定義された 4 つのグリシン残基を含む。従って、ST.26 第 7 項(b)では、この配列を配列リストに記載することを要する。

質問 3：この配列を配列表にどの様に記載すべきか？

この配列には、非準拠型の記号「z」が使用されており、その定義は開示内容から判断しなければならない（ST. 26 の「はじめに」を参照のこと）。「z」は任意のアミノ酸であると定義されているため、このアミノ酸を示すために使用される準拠型の記号は「X」である。したがって、この配列は単一の配列として示さなければならない。

GGGGX (配列番号 4)

feature key「REGION」、feature location 「>5」(>5 に対応)、およびその qualifier「note」で「The entire sequence of amino acids 1-5 can be repeated one or more times 」(アミノ酸 1~5 の全アミノ酸を、1 回またはそれ以上繰り返すことができる)と注釈を付すべきである。

第 27 項では、記号「X」は、feature table にその詳細が記載されていない限り、「A」、「R」、「N」、「D」、「C」、「Q」、「E」、「G」、「H」、「I」、「L」、「K」、「M」、「F」、「P」、「O」、「S」、「U」、「T」、「W」、「Y」または「V」のいずれか 1 つと解釈される。「X」は「任意のアミノ酸」を表しているため、feature key「VARIANT」および note qualifier (qualifier の値は「X can be any amino acid」(X は任意のアミノ酸である)と注釈を付さなければならない。

可能であれば、「X」はそれぞれに注釈を付すべきである。しかし、連続した「X」残基の領域や、入れて素性に分散した多数の「X」残基は、存在位置記述子として構文「x...y」を用いて、feature key「VARIANT」で記載しても良い。ここで、x と y は最初と最後の「X」残基の位置であり、qualifier「note」の値として「X can be any amino acid」(X は任意のアミノ酸である)を有する。

注意：上記配列の好ましい表示は、特許出願日に配列表を提供するためのものであり、提供された情報が IPO により前の出願に主題事項を追加するとみなされる可能性がある関係上、特許出願日後に提供される配列表に同表示を適用できない場合がある。

関連する ST. 26 の項：第 3 項(c)、第 7 項(b)および第 27 項

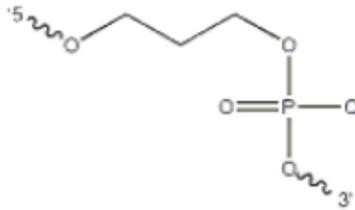
第 3 項(g) – 「ヌクレオチド」の定義

例 3(g) -1: C3 スペーサーで区切られた塩基配列

特許出願に以下の配列が記載されている。：

atgcatgcatgcncggcatgcatgc

上記の式中の「n」は C3 スペーサーであり、以下の構造を有する。



質問 1：ST. 26 では上記配列の記載を要するか？

回答：はい

この数え上げられた配列は、C3 スペーサーによって区切られた具体的に定義されたヌクレオチドで構成される 2 つのセグメントを含んでいる。

C3 スペーサーは、第 3 項(g)によればヌクレオチドではなく、準抛型の記号「n」は非準抛型の方法で使用されている（本文書の「はじめに」を参照のこと）。その結果、セグメントはそれぞれ別々の塩基配列となる。各セグメントには 10 個以上の具体的に定義されたヌクレオチドが含まれているため、その両方を配列表に記載しなければならない。

質問 3：この配列を配列表にどの様に記載すべきか？

セグメントはそれぞれ別々の配列として配列表に記載され、それぞれに配列番号が割り当てられていなければならない。

atgcatgcatgc (配列番号 5)

cggcatgcatgc (配列番号 6)

各セグメントに存在し、スペーサー-C3 と結合しているシトシンは、feature key「misc_feature」およびその qualifier「note」を用いて feature table にさらに詳しく記載すべきである。この場合の qualifier「note」の値は「free text」（フリーテキスト）で、他の核酸に結合しているスペーサーが配列上に存在していることを示し、その省略されていない完全な化学名、またはその俗称(例. C3 スペーサー)のいずれかで示すべきである。

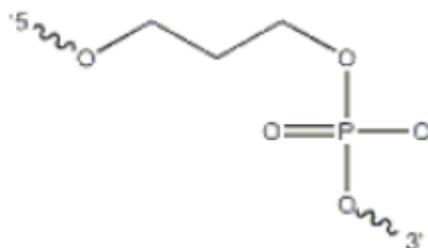
関連する ST. 26 の項：第 3 項(g)、第 7 項(a)および第 15 項

例 3(g) -2: C3 スペーサーを含む代替残基を有する塩基配列

特許出願に以下の配列が記載されている。:

atgcatgcatgcncggcatgcatgc

上記の式中の「n」は c, a, g または C3 スペーサーであり、以下の構造を有する。



質問 1: ST. 26 では上記配列の記載を要するか?

回答: はい

変数「n」で区切られた数え上げられた配列には、24 の具体的に定義された残基がある。この「n」が準拠型の方法で使用されているか、非準拠型の方法で使用されているかを判断するには、本開示の配列の説明を参照しなければならない（本文書の「はじめに」を参照のこと）。

本開示では、「n」は c, a, g または C3 スペーサーであることが示されている。「n」は、ヌクレオチドの定義を満たさない C3 スペーサーを含むと記載されていることから、非準拠型の方法で使用される準拠型の記号である。記号「n」は、「c」、「a」、「g」を含むものとしても記述されているため、ST.26 では、25 の塩基配列を配列表に記載することが求められている。C3 スペーサーで区切られた 2 つのセグメントは、25 の 塩基配列とは異なる配列であるため、2 つの 12 の塩基配列を記載しても良い。

質問 3: この配列を配列表にどの様に記載すべきか?

この例では、「n」は c, a, g または C3 スペーサーと示されている。上記の通り、C3 スペーサーはヌクレオチドではない。St. 26 の第 15 項では、記号「n」はヌクレオチド以外のものを表すために使用してはならないとされている。したがって配列表内では、「n」は、配列表において C3 スペーサーを表すことはできない。

第 15 項では、曖昧さ記号を用いるのが適切である場合でも、最も限定的な記号を使用すべきであるとしている。附属書 I、第 1 節、表 1 に従って、記号「v」を使って「c or a or g」を表しており、記号「n」よりも限定的である。

事例の変数「n」が c, a, または g である場合、最大の開示済み実施形態を含み、これにより、配列表に含まなければならない最も包括的な配列（ホン文書の「はじめに」参照）である、その残基により数え上げられる単一の配列は、以下の通りである。

atgcatgcatgcvcgcatgcatgc (配列番号 7)

また、本文書の「はじめに」で述べたように、この発明の開示または請求に不可欠な追加の配列を記載することが強く推奨される。

事例の変数「n」が C3 スペーサーである場合、配列は、変数「n」の両側に具体的に定義されたヌクレオチドの 2 つの別々のセグメント、すなわち atgcatgcatgc (配列番号 8) および、cggcatgcatgc (配列番号 9) とみなすことができる。明細書または請求項に不可欠な場合、これら 2 つの配列も、それぞれ独自の配列番号を割り当てて配列表に記載すべきである。

C3 スペーサーに結合している各セグメントのシトシンは、feature key「misc_feature」およびその qualifier「note」を用いて feature table に更に記載すべきである。この場合の qualifier「note」の値は「free text」（フリーテキスト）であり、他の核酸に結合しているスペーサーが配列上に存在していることを示し、その省略されていない完全な化学名またはその俗称（(例. C3 スペーサー)）でのいずれかによってスペーサーを識別すべきである。

注意：上記配列の好ましい表示は、特許出願日に配列表を提供するためのものであり、提供された情報が IPO により前の出願に主題事項を追加するとみなされる可能性がある関係上、特許出願日後に提供される配列表に同表示を適用できない場合がある。

関連する ST. 26 の項: 第 3 項(g)、第 7 項(a)および第 15 項

例 3(g) -3: 脱塩基部位

特許出願に以下の配列が記載されている。:

`gagcattgac-AP-taaggct`

上記の式中の AP は、脱塩基部位である。

質問 1：ST. 26 では上記配列の記載を要するか？

回答：はい

数え上げられた配列の具体的に定義された残基は、脱塩基部位で区切られている。5'位側には 10 個のヌクレオチドが含まれ、3'位側には 7 個のヌクレオチドが含まれる。第 3 項(g)(ii)(2)では、塩基配列の一部である場合、脱塩基部位を「ヌクレオチド」として定義している。その結果、本事例の脱塩基部位は、配列が配表に記載するよう求めているかどうか、およびその方法を決定する目的で、「ヌクレオチド」とみなされる。したがって、大塩基部位の領外の残基は、合計 18 個のヌクレオチドを含む 1 つの数え上げられた配列の一部であり、そのうち 17 個は具体的に定義されている。したがって、この配列は ST.26 第 7 項(a)で求められているように、配列表に 1 つの配列として含まれていなければならない。

質問 3：この配列を配列表にどの様に記載すべきか？

本配列は、配列表に以下の通り記載すべきである。

`gagcattgacntaaggct` (配列番号 10)

脱塩基部位は記号「n」で表され、さらに feature table に記載しなければならない。feature key「modified_base」およびその必須 qualifier 「mod_base」(qualifier の値は「OTHER」)を使って注釈を付す方法を推奨する。qualifier「note」には修飾塩基が脱塩基部位であることを記載しなければならない。

関連する ST. 26 の項：第 3 項(g)、第 7 項(a)および第 17 項

例 3(g) -4: 核酸類似体

特許出願に以下のグリコール核酸(GNA)の配列が記載されている。:

PO₄-tagttcattgactaaggctccccattgact-OH

上記の式中で、配列の左端が DNA 配列の 5'末端を模倣している。

質問 1：ST. 26 では上記配列の記載を要するか？

回答：はい

グリコール核酸の配列を構成する個々の残基は、ST.26 第 3 項(g)(i)(2)によればヌクレオチドとみなされる。したがって、この配列は、10 個以上の数え上げられた「具体的に定義された」ヌクレオチドを有しており、配列表への記載を要する。

質問 3：この配列を配列表にどの様に記載すべきか？

グリコール核酸の配列は、5'末端および 3'末端ではなく、3'末端および 2'末端を持つ。3'末端は、通常、末端にリン酸基を持つように表示されるが、これは DNA または RNA の 5'末端に相当する（他の核酸類似体が、DNA または RNA の 5'末端および 3'末端に別に相当する可能性があることに留意されたい）。第 11 項によると「5'末端から 3'末端への方向を模した左から右への方向で」配列リストに記載されていなければならないとされている。したがって、配列表に以下の通り記載しなければならない。

tagttcattgactaaggctccccattgact (配列番号 11)

この配列は、feature table に feature key「modified_base」およびその必須 qualifier「mod_base」(qualifier 値は「OTHER」)を使って記載しなければならない。qualifier「note」には、「glycol nucleic acids」(グリコール核酸)や「2,3-dihydroxypropyl nucleosides」(2,3-ジヒドロキシプロピルヌクレオチド) など、修飾塩基の省略されていない完全名称を記載しなければならない。

グリコール核酸の配列全体を記載するために、単一の要素「INSDFeature」を用いることも可能であり、その場合は「INSDFeature_location」の範囲は「1...30」である。

関連する ST. 26 の項：第 3 項(d)、第 3 項(g)、第 7 項(a)、第 11 項、第 16 項、第 18 項、第 65 項および第 66 項

第 3 項(k) – 「具体的に定義された」の定義

例 3(k) -1: ヌクレオチドに用いる曖昧さ記号

5' NNG KNG KNG K 3'

上の式の N および K は、国際純正・応用化学連合(IUPAC)- 国際生化学連合(IUB)共同委員会による曖昧さコードである。

質問 1: ST. 26 では上記配列の記載を要するか?

回答: いいえ

国際純正・応用化学連合(IUPAC)- 国際生化学連合(IUB)共同委員会による曖昧さコードは、附属書 I、第 1 節、表 1 で定義されたヌクレオチドに用いる記号のリストに対応している。第 3 項(k)によると、「具体的に定義された」ヌクレオチドとは、附属書 I に記載された記号「n」で表示されるヌクレオチド以外のヌクレオチドを指す。したがって、「K」および「G」は具体的に定義されたヌクレオチドであり、「N」は具体的に定義されたヌクレオチドではない。

数え上げられた配列は、10 個またはそれ以上の具体的に定義されたヌクレオチドを持たないため、ST. 26 の第 7 項(a)では、配列表に記載する必要はない。

質問 2: ST. 26 に準拠する場合、本配列の配列表への記載は許容されるか?

回答: いいえ

第 8 項では、「配列表は、10 個未満の具体的に定義されたヌクレオチドを有するいかなる配列も記載してはならない...」とある。数え上げられた配列は、10 個またはそれ以上の具体的に定義されたヌクレオチドを有していないため、配列表に記載してはならないとされている。

関連する ST. 26 の項: 第 3 項(k)、第 7 項(a)、第 8 項および、第 13 項

例 3(k) -2: 準抛型の方法と非準抛型の方法の両方で使用される曖昧さ記号「n」

出願では、次の人工配列が開示されている。5'-AATGCCGGAN-3'である。この開示は、さらに

- (i) 実施形態では、「N」は、任意のヌクレオチドである。
- (ii) 実施形態では、「N」は任意であるが、「G」にすべきである。
- (iii) 実施形態では、「N」は、「K」である。
- (iv) 実施形態では、「N」は、「C」である。

質問 1：ST. 26 では上記配列の記載を要するか？

回答：いいえ

数え上げられた配列は、9 つの具体的に定義されたヌクレオチドと「N」が含まれている。記号「N」は、準抛型の方法で使用されているかどうかは、本開示における配列の説明を参照しなければならない(本ガイダンス文書の、「はじめに」を参照のこと)。

数え上げられた配列の開示された実施形態(i)から(iv)を考慮すると、「N」の最も包括的な実施形態が「任意のヌクレオチド」であることがわかる。最も包括的な実施形態では、数え上げられた配列の「N」は準抛型の方法で使用される。

特定の実施形態では、「N」は具体的に定義された残基として記述される(すなわち、上記(iv)の「NはC」)。しかし、配列を配列表に記載しなければならないかどうかを判断する際には、最も包括的な実施形態(すなわち、「N」は任意のヌクレオチドである)のみが考慮される。したがって、評価しなければならない数え上げられた配列は、5'-AATGCCGGAN-3'である。

この分析によれば、数え上げられた配列、すなわち AATGCCGGAN は、具体的に定義された 10 個のヌクレオチドを含んでいない。したがって、ST.26 第 7 項(a)では、いくつかの実施形態において「n」も特定のヌクレオチドとして定義されているにもかかわらず、当該配列を配列リストに含めることを求めている。

質問 2：ST. 26 に準拠する場合、本配列の配列表への記載は許容されるか？

回答：いいえ

本配列「AATGCCGGAN」を配列表に記載してはならない。

ただし、「N」が具体的に定義されたヌクレオチドに置き換えられている場合は、置き換えられた配列を配列表に記載しても構わない。

質問 3：この配列を配列表にどの様に記載すべきか？

発明の重要な部分である実施形態を表す配列を記載することが強く推奨される。これらの配列を記載することで、より完全な検索が可能になり、特許を求める主題事項を公に示すことができる。

上述の例では、以下の 3 つの追加の配列を、それぞれ独自の配列番号を付けて配列表に記載することが強く推奨される。

aatgccggag (配列番号 12)

aatgccggak (配列番号 13)

aatgccggac (配列番号 14)

上述の配列 3 つすべてを記載しない場合、「n」の部位を置き換えるヌクレオチドは、置き換えた旨を記述するために注釈を付すべきである。例えば、上記の配列番号 12 のみを配列表に記載する場合、feature key「misc_difference」および feature location 「10」を使用し、2 つの qualifier「replace」(片方の値を「g」、もう一方の値を「c」)を併用すべきである。

注意：上記配列の好ましい表示は、特許出願日に配列表を提供するためのものであり、提供された情報が IPO により前の出願に主題事項を追加するとみなされる可能性がある関係上、特許出願日後に提供される配列表に同表示を適用できない場合がある。

関連する ST. 26 の項：第 3 項(k)、第 7 項(a)、第 8 項および、第 13 項

例 3(k) -3: 非準拠型の方法で使用される曖昧さ記号「n」

出願では、次の配列が記載されている。5'-aatgttggan-3'である。

ここで、n は c である。

質問 1：ST. 26 では上記配列の記載を要するか？

回答：はい

第 3 項(k)によると、「具体的に定義された」ヌクレオチドとは、附属書 I、第 1 節、表 1 に記載されている記号「n」で表されるヌクレオチド以外のヌクレオチドのことである。

この例で、「n」は「c」のみを表すために非準拠型の方法で使用されている。本開示は、「n」が「任意のヌクレオチド」を表すために準拠型の方法で使用されることを示していない。したがって、配列は、同等の準拠型の記号、すなわち「c」が配列に使用されたかのように解釈しなければならない（本文書の「はじめに」を参照のこと）。したがって、考慮しなければならない数え上げられた配列とは、以下の通りとなる。

5'-aatgttggac-3'

この配列は、10 個の具体的に定義されたヌクレオチドを有しており、ST. 26 第 7 項(a)により、配列表に記載することを要する。

質問 3：この配列を配列表にどの様に記載すべきか？

aatgttggac (配列番号 15)のように配列表に記載されている必要がある。

関連する ST. 26 の項：第 3 項(k)および第 7 項(a)

例 3(k) -4: 「n」以外の曖昧さ記号が「具体的に定義された」残基を示している場合

出願には、以下の配列が記載されている。

5' NNG KNG KNG KAG VCR 3'

ここで、N、K、V および R は、国際純正・応用化学連合(IUPAC)- 国際生化学連合(IUB)共同委員会による曖昧さコードである。

質問 1：ST. 26 では上記配列の記載を要するか？

回答：はい

国際純正・応用化学連合(IUPAC)- 国際生化学連合(IUB)共同委員会による曖昧さコードは、附属書 I、第 1 節、表 1 に定義されているヌクレオチド記号の表に対応している。第 3 項(k)によると、「具体的に定義された」ヌクレオチドとは、附属書 I、第 1 節、表 1 に記載されている記号「n」で表されるヌクレオチド以外のヌクレオチドを指す。したがって、「K」、「V」および「R」は「具体的に定義された」ヌクレオチドである。

この配列は、11 個の数え上げられた「具体的に定義された」ヌクレオチドを有し、ST. 26 第 7 項(a)により、配列表への記載を要する。

質問 3：この配列を配列表にどの様に記載すべきか？

以下の通り配列表に記載しなければならない。

nngkngkngkagvcr (配列番号 16)

関連する ST. 26 の項：第 3 項(k)、第 7 項(a)および第 15 項

例 3 (k) -5: 非準拠型の方法で使用される曖昧さ記号「Xaa」

特許出願に、以下の配列が記載されている。

Xaa-Tyr-Glu-Xaa-Xaa-Xaa-Leu

上記の式で、1 位の記号「Xaa」は任意のアミノ酸であり、4 位の記号「Xaa」は Lys、5 位の記号「Xaa」は Gly、6 位の「Xaa」はロイシンまたはイソロイシンである。

質問 1：ST. 26 では上記配列の記載を要するか？

回答：はい

上記の式の数え上げられたペプチドは、2 位、3 位および 7 位に 3 つの具体的に定義されたアミノ酸が示されている。1 位のアミノ酸は、任意のアミノ酸を表す準拠型の略語、すなわち「Xaa」で表される。しかし 4 位、5 位および 6 位のアミノ酸は、準拠型の略語が非準拠型の方法で使用されている（本文書の「はじめに」を参照のこと）。したがって、これらの位置に存在する「Xaa」の定義を確認するために、開示されている配列の説明を参照する。4 位から 6 位の「Xaa」は特定のアミノ酸として示されているため、配列は、同等の準拠型の略語、すなわち、Lys、Gly および (Leu または Ile) が配列に使用されているかのように解釈しなければならない。その結果、本配列は 4 つまたはそれ以上の具体的に定義されたアミノ酸を含んでおり、ST. 26 第 7 項(b)で求められているように、配列表に記載しなければならない。

質問 3：この配列を配列表にどの様に記載すべきか？

この配列で、準拠型の略語である「Xaa」が非準拠型の方法で使用されている。そのため、4 位、5 位、6 位の「Xaa」の定義を確認するためには、開示されている配列の説明を参照する必要がある。この説明では、4 位のリジン、5 位のグリシン、6 位のロイシンまたはイソロイシンが「Xaa」と定義されている。これらのアミノ酸を表す準拠型の記号は、それぞれ K、G、J である。したがって、配列表では以下のように記載すべきである。

XYEKGJL (配列番号 17)

第 27 項によれば、上記の式中の記号「X」は、feature table に詳細な情報がある場合を除き、「A」、「R」、「N」、「D」、「C」、「Q」、「E」、「G」、「H」、「I」、「L」、「K」、「M」、「F」、「P」、「O」、「S」、「U」、「T」、「W」、「Y」または「V」のいずれか 1 つであると解釈する。配列番号 17 の 1 位の「X」は「任意のアミノ酸」を表しているため、feature key「VARIANT」およびその qualifier「note」(qualifier の値として「X can be any amino acid」(「X」は任意のアミノ酸であり得る) との注釈を付さなければならない。

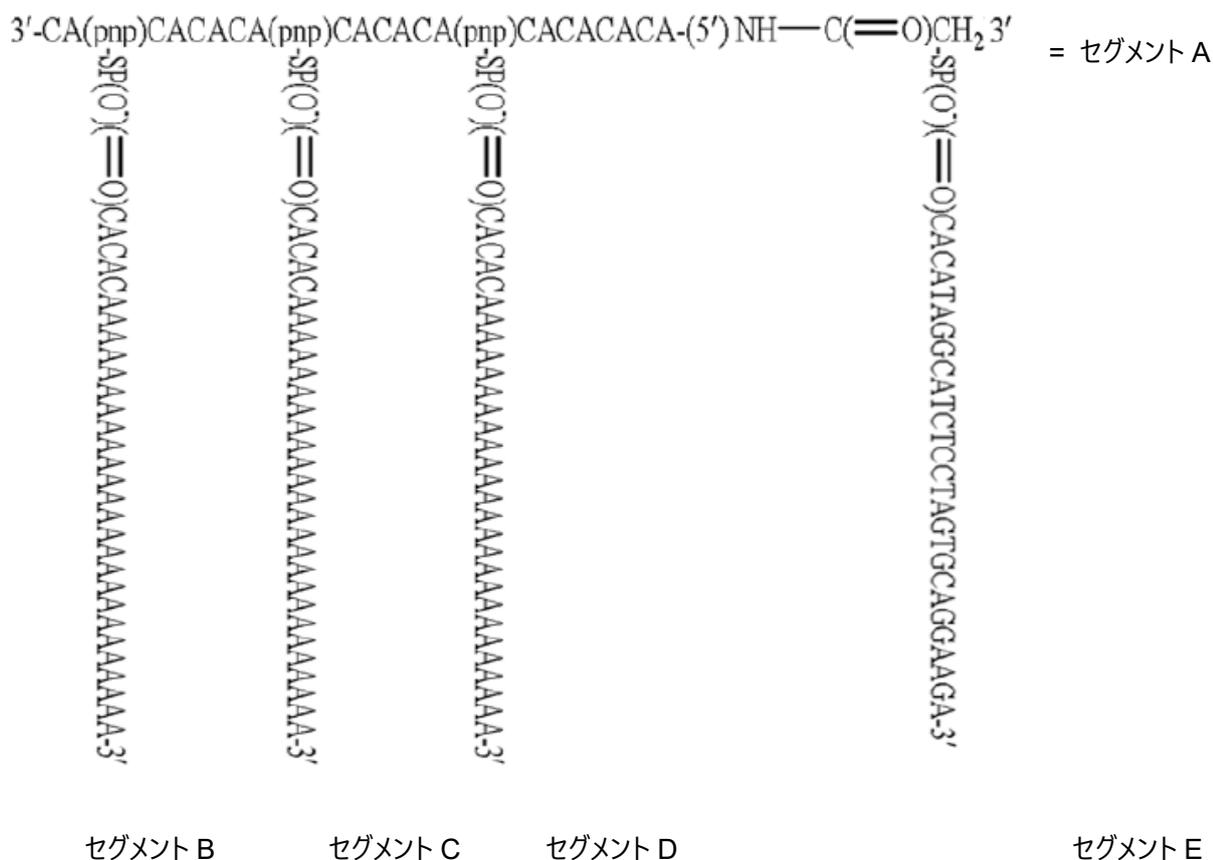
可能であれば、「X」それぞれに対して個別に注釈を付すべきである。しかし、連続した「X」残基が複数存在する領域や、配列上に分散した多数の「X」残基は、存在位置記述子として構文「x...y」(x と y は先頭と末尾の「X」残基) を用いて、feature key「VARIANT」およびその qualifier「note」(qualifier 値は「X can be any amino acid」(「X」は任意のアミノ酸であり得る) を用いてまとめて記載しても良い。

関連する ST. 26 の項：第 3 項(k)、第 7 項(b)、第 26 項および第 27 項

第 7 項(a) - 配列表に必要な塩基配列

例 7(a) -1: 分岐塩基配列

明細書に、以下の分枝連鎖の配列が記載されている。



上記の式中の記号「pnp」は、プロモアセチルアミノ官能基を含む結合またはモノマーである。

3'-CA(pnp)CACACA(pnp)CACACA(pnp)CACACACA-(5')NH—C(=O)CH₂ 3' はセグメント A である。

SP(O)-(=O)CACACAAAAAAAAAAAAAAAAAAAAAAAAA 3' はセグメント B、C および D である。さらに、

SP(O)-(=O)CACATAGGCATCTCCTAGTGCAGGAAGA 3' はセグメント E である。

質問 1： ST. 26 では上記配列の記載を要するか？

回答： はい - 4 つの縦方向のセグメント B~E は配列表に記載しなければならない。

いいえ - 横方向のセグメント A は配列表に記載してはならない。

上記の図は、5 つの線形のセグメント（横方向のセグメント A および 4 つの縦方向のセグメント B~E）を含む「櫛型」分枝核酸配列の例である。

第 7 項(a)によると、10 個またはそれ以上の具体的に定義されたヌクレオチドを含み、隣接するヌクレオチドが 3'-位から 5'-位に結合している分岐した塩基配列は、配列表に記載しなければならない。

4 本の縦方向のセグメント B~E は、それぞれが 10 個以上の具体的に定義されたヌクレオチドを含み、隣接するヌクレオチドが 3'-位から 5'-位に結合しているため、それぞれが配列表に記載される必要がある。

横方向のセグメント A では、塩基配列の線形の領域が非塩基配列の部位「pnp」と結合しており、「pnp」を介して結合しているこれらの線形の領域のそれぞれが含む具体的に定義されたヌクレオチドは 10 個未満である。したがって、セグメント A のどの領域も、隣接するヌクレオチドが 3'-位から 5'-位に結合している 10 個以上の具体的に定義されたヌクレオチドを含んでいないため、それらは、ST. 26 第 7 項(a)によって配列表への記載を要しない。

質問 2： ST. 26 に準拠する場合、本配列の配列表への記載は許容されるか？

回答： いいえ

第 8 項では、「配列表は、10 個未満の具体的に定義されたヌクレオチドを有するいかなる配列も含んではならない...」とある。セグメント A のどの領域も、隣接するヌクレオチドが 3'-位から 5'-位に結合している 10 個以上の具体的に定義されたヌクレオチドを含んでいない。したがって、個別の配列番号を持つ独立した配列として配列表に記載してはならない。

ただし、セグメント B、C、D、E はセグメント A と結合していることを示すために注釈を付しても良い。

質問 3： この配列を配列表にどの様に記載すべきか？

セグメント B、C、D は同一の配列であり、配列表には以下の 1 つの配列として記載しなければならない。

cacacaaaaaaaaaaaaaaaaaaaaaaaaa (配列番号 18)

配列の先頭の記号「c」は、feature key「misc_feature」およびその qualifier「note」を用いて、例えば、「This sequence is one of four branches of a branched polynucleotide」（この配列は、分枝連鎖のポリヌクレオチドの 4 つの連鎖の 1 つである）のように、詳細な情報を記載すべきである。

セグメント E は、以下の単一配列として配列表に記載しなければならない。

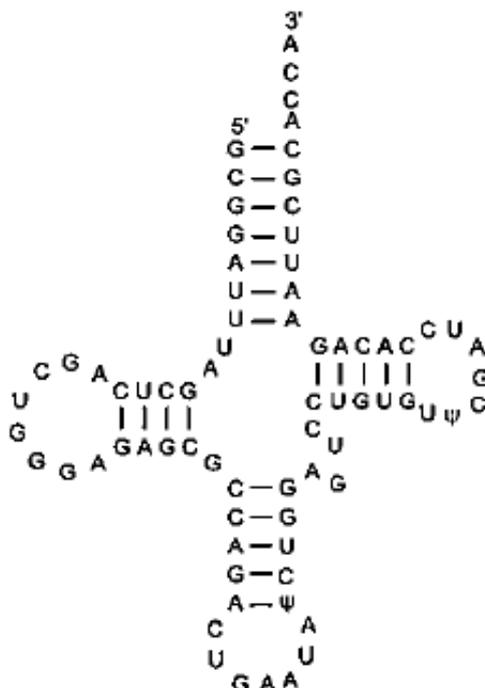
cacataggcatctcctagtcaggaaga (配列番号 19)

配列の先頭の記号「c」は、feature key「misc_feature」およびその qualifier「note」を用いて、例えば、「This sequence is one of four branches of a branched polynucleotide」（この配列は、分枝連鎖のポリヌクレオチドの 4 つの連鎖の 1 つである）のように、詳細な情報を記載すべきである。

関連する ST. 26 の項： 第 7 項(a)、第 8 項、第 11 項、第 13 項および第 17 項

例 7(a) -2: 二次構造を有する線形の塩基配列

特許出願に、以下の配列が記載されている。



上の式のΨはシュドウリジンである。

質問 1 : ST. 26 では上記配列の記載を要するか?

回答 : はい

この塩基配列には、73 個の数え上げられた具体的に定義されたヌクレオチドが含まれている。したがって、この例は 10 個またはそれ以上の「具体的に定義された」ヌクレオチドを有しており、ST. 26 第 7 項(a)で求められているように、配列表に記載しなければならない。

質問 3 : この配列を配列表にどの様に記載すべきか?

明細書の説明では、記号「Ψ」はシュドウリジンと同等であることがわかる。シュドウリジンを示すために使用できる唯一の準拠型の記号は「n」である。したがって、「Ψ」は準拠型の記号「n」を表すために用いられる非準拠型の記号である（本文書の「はじめに」を参照のこと）。したがって、この配列は、2 つの記号「Ψ」の代わりに 2 つの記号「n」があるものと解釈しなければならない。

記号「u」は、配列表の RNA 分子のウラシルを表すために使用してはならない。第 14 項によると、記号「t」は RNA を構成するウラシルであると解釈する。この配列は以下のように記載しなければならない。

gcggttagctcagctgggagagcgccagactgaatanctggagtcctgtgtncgatccacagaattcgacca (配列番号 20)

必須 feature key「source」の qualifier「mol_type」の値は「tRNA」である。また、詳細な情報は feature key「tRNA」および任意の適切な qualifier のいずれかを用いて記載しても良い。

記号「n」で示されている残基の詳細な情報は、feature key「modified_base」およびその必須 qualifier「mod_base」を用いた feature table に、qualifier 値としてシュドウリジンを示す略語「p」を用いて記載しなければならない(附属書 1、表 2 を参照のこと)。

関連する ST. 26 の項 : 第 7 項(a)、第 11 項、第 13 項、第 14 項、第 17 項、第 62 項、第 84 項、附属書 1 第 2 節および第 5 節、feature key5.43

例 7(a) -3: 非準拠型の方法で使用されるヌクレオチドの曖昧さ記号

特許出願に、以下の配列が記載されている。

5' GATC-MDR-MDR-MDR-MDR-GTAC 3'

明細書中の配列の説明では、次の記載がある。「DR 要素」は配列 5' ATCAGCCAT 3'で構成されている。DR 要素の変異、すなわち MDR は、中央の 5 つのヌクレオチド「CAGCC」が「TTTTT」に変異した DR 要素である。

質問 1：ST. 26 では上記配列の記載を要するか？

回答： はい

数え上げられた配列では、記号「MDR」が用いられている。配列に使用されている記号が、準拠型の記号、すなわち、附属書 1、第 3 節、表 3 に記載されている記号を意図しているのか、それとも非準拠型の記号を意図しているのかが不明な場合は、明細書の配列に関する説明を参照して判断しなければならない（本文書の「はじめに」を参照のこと）。表 3 によれば、「MDR」は、3 つの準拠型の記号(m = a または c, d = a または g または t/u, r = g または a)として解釈されるか、あるいは、他の構造を表す略記号として解釈される。

明細書中の説明を参照すると、MDR 要素は 5' ATTTTTTAT 3'と同等であることがわかる。「MDR」の文字は、非準拠型の方法で使用される準拠型の記号と考えられ、したがって、配列は、同等の準拠型の記号を使用して開示されているかのように解釈されなければならない。したがって、配列表への記載を要すると思われる数え上げられた配列は以下のとおりである。

5' GATC ATTTTTTAT ATTTTTTAT ATTTTTTAT ATTTTTTAT GTAC 3'

数え上げられた配列は、44 個の具体的に定義されたヌクレオチドを有し、ST. 26 第 7 項(a)により、配列表への記載を要する。

質問 3：この配列を配列表にどの様に記載すべきか？

配列は以下のように、配列表に記載しなければならない。

gatcatttttatatttttatatttttatatttttatgtac (配列番号 21)

関連する ST. 26 の項： 第 7 項(a)および第 13 項

例 7(a) -4: 非準拠型の方法で使用されるヌクレオチドの曖昧さ記号

特許出願に、以下の配列が記載されている。

5' ATTC-N-N-N-N-GTAC 3'

明細書における配列の説明では、次の通り示されている。記号「N」は配列 5' ATACGCACT 3' で構成されている。

質問 1： ST. 26 では上記配列の記載を要するか？

回答： はい

数え上げられた配列では、記号「N」が用いられている。「N」が準拠型の方法で使用されているか、非準拠型の方法で使用されているかを判断するには、明細書における配列の説明を参照しなければならない（本文書の「はじめに」を参照のこと）。

明細書を参照すると、「N」は、5' ATACGCACT 3' と同等であることがわかる。このように、「N」は、準拠型の記号を非準拠型の方法で使用したものである。したがって、この配列は、同等の準拠型の記号を使用して表示されたかのように解釈しなければならない。

5' ATTC-ATACGCACT-ATACGCACT-ATACGCACT-ATACGCACT-GTAC 3'

質問 3： この配列を配列表にどの様に記載すべきか？

配列は以下のように、配列表に記載しなければならない。

atctacgcactatacgcactatacgcactatacgcactgtac (配列番号 22)

関連する ST. 26 の項： 第 7 項(a)および第 13 項

例 7(a) -5: 非準拠型のヌクレオチドの記号

特許出願に、以下の配列が記載されている。

5' GATC-β-β-β-β-GTAC 3'

明細書における配列の説明では、次の通り示されている。記号「β」は配列 5'ATACGCACT 3'で構成されている。

質問 1：ST. 26 では上記配列の記載を要するか？

回答： はい

数え上げられた配列では、非準拠型の記号「β」が用いられている。この「β」の意味については、明細書における配列の説明を参照しなければならない（本文書の「はじめに」を参照のこと）。

明細書における配列の説明を参照すると、「β」は 5' ATACGCACT 3'と同等であることがわかる。このように、「β」は、具体的に定義された 9 つの準拠型の記号の配列を表すために使用される非準拠型の記号である。したがって、この配列は、同等の準拠型の記号を用いて表示されているかのように解釈しなければならない。

5' GATC-ATACGCACT-ATACGCACT-ATACGCACT-ATACGCACT-GTAC 3'

数え上げられた配列は、44 個の具体的に定義されたヌクレオチドを有し、ST. 26 第 7 項(a)により、配列表への記載を要する。

質問 3：この配列を配列表にどの様に記載すべきか？

配列は以下のように、配列表に記載しなければならない。

gatcatacgcactatacgcactatacgcactatacgcactgtac (配列番号 23)

関連する ST. 26 の項： 第 7 項(a)および第 13 項

例 7(a) -6: 非準拠型のヌクレオチドの記号

特許出願に、以下の配列が記載されている。

5' GATC-β-β-β-β-GTAC 3'

明細書における配列の説明では、次の通り示されている。記号「β」はアデニン、イノシンまたはシュドウリジンと同等である。

質問 1：ST. 26 では上記配列の記載を要するか？

回答： いいえ

数え上げられた配列では、記号「β」が用いられている。この「β」の意味については、明細書における配列の説明を参照しなければならない（本文書の「はじめに」を参照のこと）。

明細書中における配列の説明を参照すると、「β」は、アデニン、イノシンまたはシュドウリジンと同等であることがわかる。アデニン、イノシンまたはシュドウリジンを示すために用いることができる唯一の準拠型の記号は「n」であり、したがって、「β」は準拠型の記号「n」を表すために用いられる非準拠型の記号である。したがって、この配列は、4つの「β」の代わりに、4つの「n」（以下、「N」と表記）を有すると解釈しなければならない。

5' GATC-N-N-N-N-GTAC 3'

この数え上げられた配列は、具体的に定義された8つのヌクレオチドしかなく、ST. 26 第7項(a)では、配列表への記載を要しない。

質問 2：ST. 26 に準拠する場合、本配列の配列表への記載は許容されるか？

回答： いいえ

数え上げられた配列である 5' GATC-N-N-N-N-GTAC 3' は配列表に記載してはならない。

ただし、開示されている代替りの配列は、記号「n」のうち少なくとも2つがアデニンと置き換えられ、その結果、少なくとも10個以上の具体的に定義されたヌクレオチドを有する配列となっていれば、配列表に記載しても良い。

質問 3：この配列を配列表にどの様に記載すべきか？

許容される表示の一つとしては、以下の通りである。

gatcaaaagtac (配列番号 24)

上記の例では、記号「β」を置換する4つのアデニンヌクレオチドは、これらの位置に存在するアデニンヌクレオチドが、イノシンまたはシュドウリジンと置き換わる可能性があることを注釈を付すべきである。

feature key「misc_difference」に、feature location「5-8」および qualifier「note」の値、例えば、「A nucleotide in any of position 5-8 may be replaced with inosine or pseudouridine」（5'-位から8'-位の位置にあるいずれかのヌクレオチドは、イノシンまたはシュドウリジンと置き換わる可能性がある）を用いて記載すべきである。これらの選択肢は、修飾塩基であるため、feature key「modified_base」および qualifier「mod_base」が必要となる。この場合の qualifier「modified_base」の値は「OTHER」に qualifier「note」と「i or p」の値を加えたものであり得る。

その他にも様々な組み合わせが可能である。

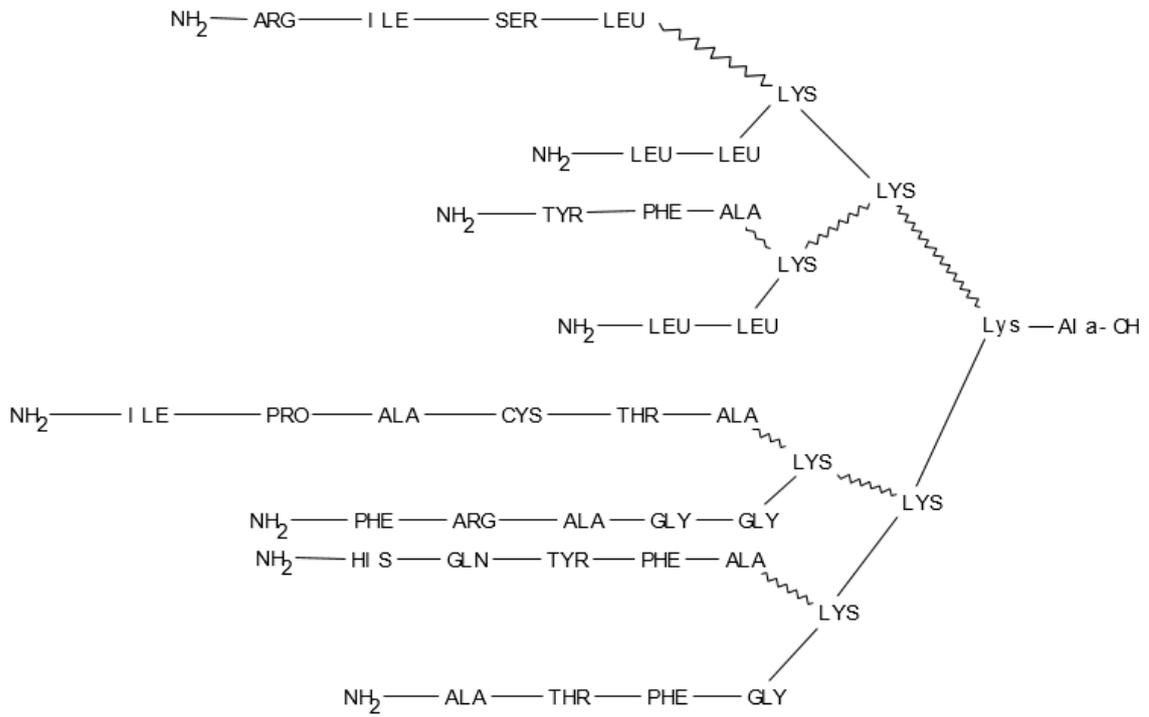
注意：上記配列の好ましい表示は、特許出願日に配列表を提供するためのものであり、提供された情報が IPO により前の出願に主題事項を追加するとみなされる可能性がある関係上、特許出願日後に提供される配列表に同表示を適用できない場合がある。

関連する ST. 26 の項：第7項(a)、第8項、第13項および第17項
第7項(b) - 配列表に必要なアミノ酸配列

例 7(b) -1: 4つまたはそれ以上の具体的に定義されたアミノ酸

例 7(b) -2: 分岐アミノ酸配列

この出願では、リシン残基を足場とした8つの分枝を形成し、そこに複数の直鎖状ペプチド鎖を結合させた分岐した配列を記載している。リシンは二塩基アミノ酸で、2箇所までペプチド結合している。このペプチドの配列は以下のとおりである。

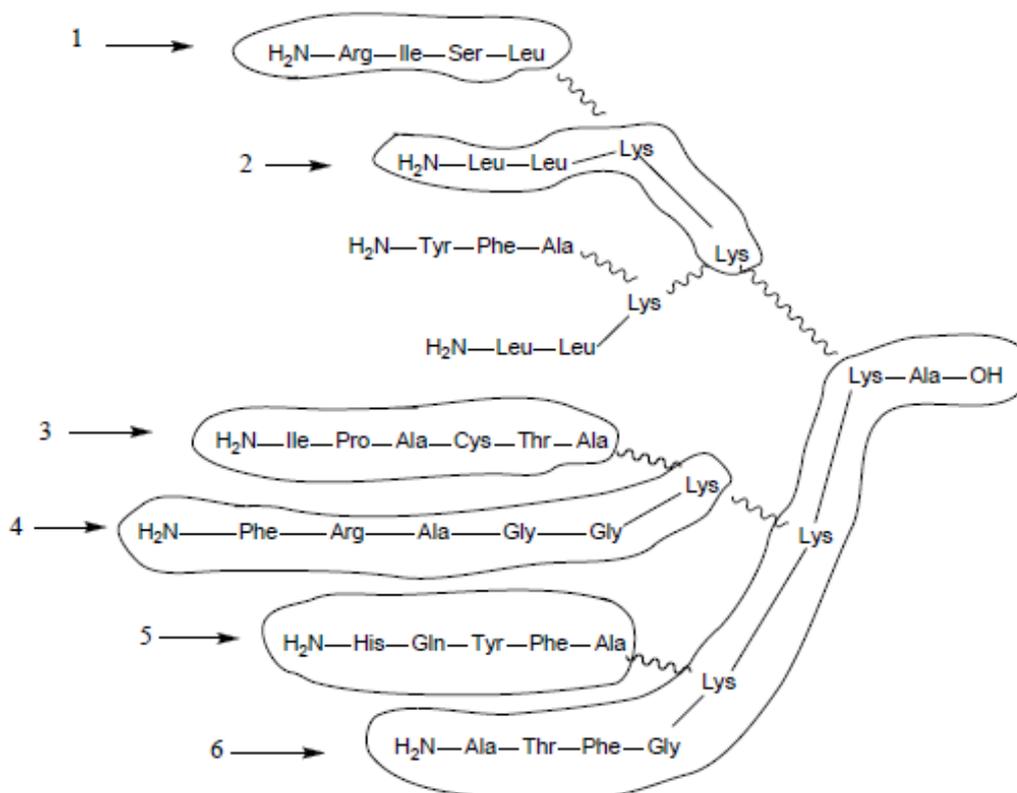


上記の分枝連鎖のペプチドにおいて、____で表されているリシンと他のアミノ酸との結合は、リシンの末端アミンと結合したアミノ酸のC末端との間のアミド結合を表している。波線(~~~~~)で表されている結合は、リシンの側鎖アミンと結合したアミノ酸のC末端とのアミド結合を表している。

質問 1：ST. 26 では上記配列の記載を要するか？

回答：はい

この例では、リシン残基が足場として使用されている分枝した配列が表示されている。第 7 項(b)では、4 つ以上の具体的に定義されたアミノ酸を含む、非分枝領域または線形領域を配列表に記載することが求められている。上記の例では、4 つ以上の例では、4 つ以上の具体的に定義されたアミノ酸を含む分枝連鎖のペプチドの線形領域を囲んである。



ST. 26 第 7 項(b)では、上記のペプチド 1～6 を配列表に記載することを要する。配列表への記載が求められていないペプチドは以下の通りである。

YFA

LLK

質問 2：ST. 26 に準拠する場合、本配列の配列表への記載は許容されるか？

回答：いいえ

第 8 項によれば、配列表には、4 つ未満の具体的に定義されたアミノ酸を有する配列を記載してはならない。

ペプチド配列 YFA および LLK は、それぞれ 3 つの具体的に定義されたアミノ酸しか有しておらず、個別の配列番号を有する別の配列として、配列表に記載してはならない。

質問 3：この配列を配列表にどの様に記載すべきか？

ペプチド 1～6 は、個別に配列番号で表さなければならない。

RISL (配列番号 26)

LLKK (配列番号 27)

IPACTA (配列番号 28)

FRAGGK (配列番号 29)

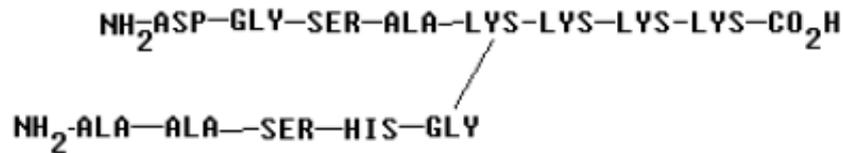
HQYFA (配列番号 30)
ATFGKKKA (配列番号 31)

架橋は、feature key「SITE」および必須 qualifier「note」を用いて、例えば「This sequence is one part of a branched amino acid sequence」（この配列は、分枝連鎖のアミノ酸配列の一部である）という値で表記することができる。ST. 26 の第 30 項によれば、配列番号 27、29、31 は、各リシンについて、feature key「SITE」および、リシンの側鎖が他の配列とアミド結合を介して結合していることを記述する qualifier「note」を用いて、修飾アミノ酸であることを示す注釈を付さなければならない。配列番号 26、28、30 のそれぞれには、feature key「SITE」およびその qualifier「note」を用いて C 末端のアミノ酸が別の配列に結合していることを示す注釈を付すべきである。

関連する ST. 26 の項：第 7 項(b)、第 8 項、第 26 項、第 29 項、第 30 項および第 31 項

例 7(b) -3: 分岐アミノ酸配列

以下、ペプチド配列を示す。



下段の配列の末端に位置するグリシン残基と、上段の配列のリシンの間の結合は、グリシンのカルボキシ末端とリシンのアミノ末端側鎖の間のアミド結合を介して行われる。

質問 1：ST. 26 では上記配列の記載を要するか？

回答：はい

4 つ以上の具体的に定義されたアミノ酸を含む、配列の非分枝領域または線形領域は、配列表に記載しなければならない。上記の例では、4 つ以上のアミノ酸を有する分枝連鎖のペプチドの線形領域を囲んである。



ST. 26 第 7 項(b)では、配列表に配列 1 および 2 を記載することを要する。

質問 3：この配列を配列表にどの様に記載すべきか？

配列 1 および 2 は、個別の配列番号で示されなければならない。

DGSAK KKK (配列番号 32)

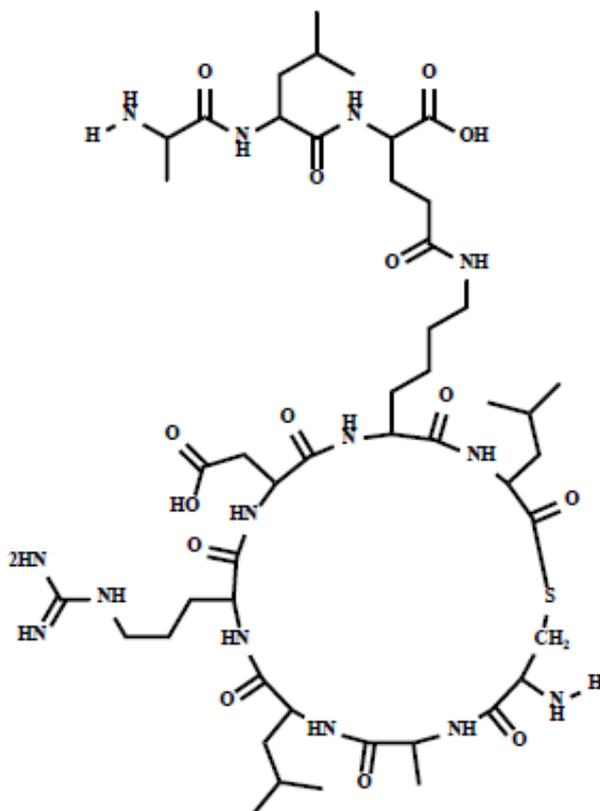
AASHG (配列番号 33)

配列 DGSAK KKK は、feature key「SITE」および、リシンの側鎖がアミド結合を介して他の配列に結合していることを示す qualifier「note」を用いて、5 位のリシンが修飾アミノ酸であることを示す注釈を付さなければならない。配列 AASHG は、feature key「SITE」および qualifier「note」を用いて、5 位のグリシンが他の配列に結合していることを示す注釈を付すべきである。

関連する ST. 26 の項：第 7 項(b)、第 26 項、第 29 項、第 30 項および第 31 項

例 7(b) -4: 分岐アミノ酸配列を含む環状ペプチド

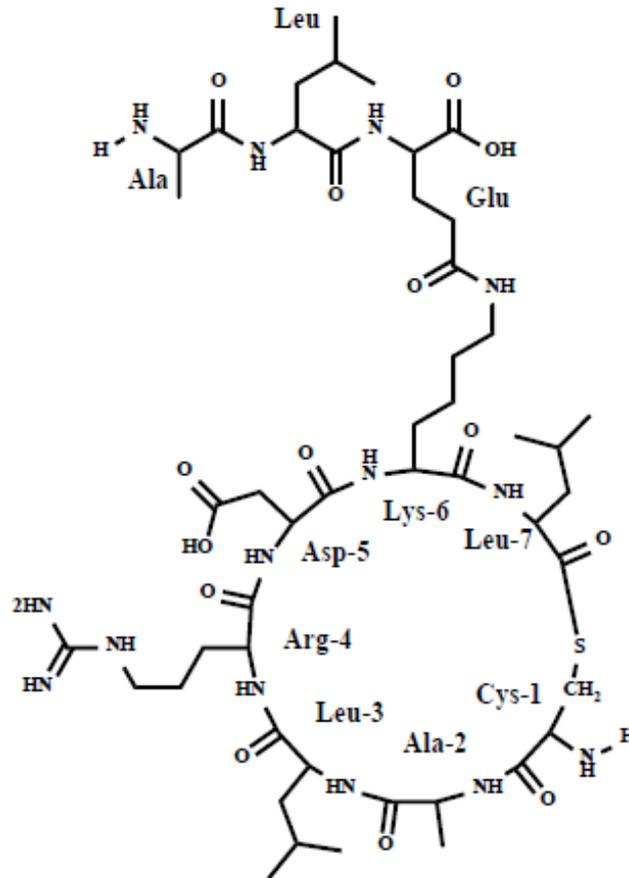
特許出願には、以下の構造が記載されている。



環状構造のシステインとロイシンは、システインの側鎖およびロイシンのカルボキシ末端を介して結合している。

質問 1：ST. 26 では上記配列の記載を要するか？

この構造は、分枝連鎖した環状のアミノ酸配列で、以下のアミノ酸が含まれている。



システインの側鎖およびロイシンのカルボキシ末端が環状構造を構成しているため、環状ペプチドの N 末端は Cys-1 に位置している。

回答：はい - ペプチド鎖の環状領域

ST.26 第 7 項(b)では、4 つ以上の具体的に定義されたアミノ酸を含む分枝配列の線形領域で、アミノ酸が単一のペプチド骨格を形成しているものは、配列表に記載しなければならないとしている。上記の例では、分枝したペプチドの環状領域には 4 つ以上のアミノ酸が含まれているため、配列表に記載しなければならない。

回答：いいえ - ペプチドのトリペプチド分枝

Ala-Leu-Glu のトリペプチド分枝は、配列表に記載する必要はない。

質問 2：ST. 26 に準拠する場合、本配列の配列表への記載は許容されるか？

回答：いいえ

第 8 項によると、配列表には、具体的に定義されたアミノ酸の数が 4 つ未満の配列を記載してはならないとされている。

トリペプチド分枝は、3 つの具体的に定義されたアミノ酸しか含まれていないため、個別の配列番号を持つ独立した配列として、配列表に記載してはならない。

質問 3：この配列を配列表にどの様に記載すべきか？

この例では環状に配置されたペプチドが示されているが、第 25 項に示されているように、環はペプチド結合のアミノ酸残基だけで構成されているわけではない。アミノ酸配列の環化は、システイン(Cys)の側鎖とロイシン(Leu)のカルボキシ末端を介して行われるた

め、ペプチドの環状領域内では、システインは先頭の位置番号が割り当てられなければならない。したがって、この配列は以下の通り表さなければならない。

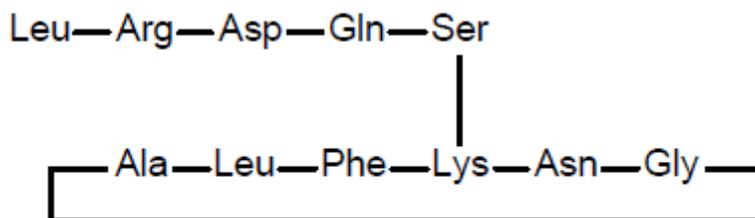
CALRDKL (配列番号 89)

上の図に示されているように、アミノ酸配列は、システイン側鎖とロイシンのカルボキシ末端との間のチオエステル結合を介して環状になっている。ロイシンとの間に鎖内結合を形成する修飾されたシステインを記載するために、feature key「SITE」を使用しなければならない。feature location の要素は、架橋されたアミノ酸の残基番号を「x..y」形式、すなわち「1...7」で表している。必須 qualifier「note」は結合の性質を示すべきであり、例えば「システイン・ロイシン・チオエステル (Cys-Leu)」とすることで、Cys-1 と Leu-7 がチオエステル結合を介して結合していることを指定する。さらに、6 位のロイシンは修飾されていることを示すために、feature key「SITE」と必須 qualifier「note」を用いて注釈を付さなければならない、qualifier 値はロイシン側鎖がトリペプチド ALE を結合していることを表している。

関連する ST. 26 の項：第 7 項(b)、第 8 項、第 25 項、第 26 項、第 29 項、第 30 項、第 31 項、第 66 項(c)および第 70 項。

例 7(b) -5: 分岐アミノ酸配列を含む環状ペプチド

特許出願に、以下の分枝した環状配列が記載されている。

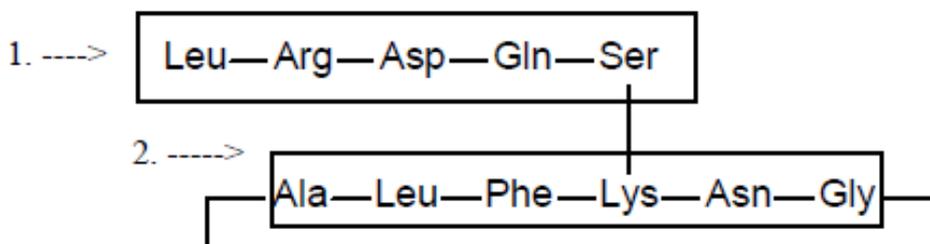


セリン(Ser)およびリシン(Lys)は、セリンのカルボキシ末端およびリシンの側鎖の間のアミド結合を介して結合している。

質問 1：ST. 26 では上記配列の記載を要するか？

回答：はい

第 7 項(b)では、4 つ以上の具体的に定義されたアミノ酸を含み、分枝した配列の線形領域として表現できる配列を、配列表に記載することを要する。上記の例では、ペプチドは、アミノ酸がペプチド結合で結合している環状領域および、環状領域に存在するリシン(Lys)の側鎖に結合している分枝領域を含んでいる。この分枝したペプチドのうち、線形状に表すことができ、かつ 4 つ以上の具体的に定義されたアミノ酸を含む領域は、以下の通りである。



ST. 26 では、この環状分枝ペプチドの配列 1 および 2 を、それぞれの配列番号を割り当てて配列表に記載することを要する。

質問 3：この配列を配列表にどの様に記載すべきか？

上記の配列 1 は以下の通り示さなければならない。

LRDQS (配列番号 90)

配列 1 は、feature key「SITE」および qualifier「note」を用いて、5 位のセリンが他の配列のリシン(Lys)の側鎖との間のアミド結合を介して他の配列に結合していることを示す注釈を付すことができる。

配列 2 は、環状ペプチドである。第 25 項は、アミノ酸配列が環状構造で、その環は、アミノ末端およびカルボキシ末端を持たない場合、出願人は位置番号 1 のアミノ酸残基を選択しなければならないことを示している。したがって、この配列は以下の通り示しても良い。

ALFKNG (配列番号 91)

また、配列内の任意の他のアミノ酸を残基番号 1 として指定することも可能である。配列 ALFKNG は、feature key「SITE」および qualifier「note」を用いて、残基番号 4 のリシン(Lys)の側鎖がアミド結合を介して他の配列に結合していることを説明しなければならない。この側鎖の結合はリシン(Lys)を修飾するものであり、ST. 26 第 30 項によれば、修飾を受けたアミノ酸は feature table に詳細に記載しなければならない。さらに、ペプチド ALFKNG が環状であることを示すために、feature key「REGION」および qualifier「note」を用いて記載すべきである。

関連する ST. 26 の項：第 7 項(b)、第 25 項、第 26 項、第 30 項および第 31 項

第 11 項(a) - 二本鎖の塩基配列 - 完全に補完しあう塩基配列

例 11(a) -1: 二本鎖の塩基配列 - 長さが同じもの

特許出願には、以下の二本鎖の DNA 配列が記載されている。

3'-CCGGTTAACGCTA-5'

5'-GGCCAATTGCGAT-3'

質問 1：ST. 26 では上記配列の記載を要するか？

回答：はい

それぞれ数え上げられた塩基配列は、10 個以上の具体的に定義されたヌクレオチドを有する。この二本鎖の塩基配列の 2 本の鎖は互いに完全に相補的であるため、少なくとも 1 本の鎖を配列表に記載しなければならない。

質問 2：ST. 26 に準拠する場合、本配列の配列表への記載は許容されるか？

回答：はい

配列表には、片方の鎖の配列のみを記載しなければならないが、両鎖の配列にそれぞれ個別の配列番号を割り当てて記載しても良い。

質問 3：この配列を配列表にどの様に記載すべきか？

二本鎖の DNA 配列は、単一配列として、または 2 つの別々の配列として表現しなければならない。配列表に記載する各配列は、5 位から 3 位方向に表示しなければならず、それぞれ個別の識別番号を割り当てなければならない。

atcgcaattggcc (top strand) (配列番号 34)

および／または、

ggccaattgcat (bottom strand) (配列番号 35)

関連する ST. 26 の項：第 7 項(a)、第 11 項(a)および第 13 項

第 11 項(b) - 二本鎖の塩基配列 - 完全に補完しあう塩基配列ではない塩基配列

例 11(b) -1: 二本鎖の塩基配列 - 長さが異なるもの

特許出願には、以下の配列図および説明が記載されている。

```

5'-tagttcattgactaaggctccccattgactaaggcgactagcattgactaaggcaagc-3'
      |||
      gggtactgantccgc
  
```

ペプチド核酸(PNA)プローブ (下の鎖)が結合したヒト遺伝子 ABC1 プロモーター領域 (上の鎖)であり、PNA プローブの配列上の記号「n」は、5 位のニトロインドールから 3 位のニトロインドールからなる群から選択されるユニバーサル PNA 塩基である。

質問 1 : ST. 26 では上記配列の記載を要するか?

回答 : はい - ABC1 プロモーター領域 (上の鎖)

上の鎖は 10 個以上の数え上げられた「具体的に定義された」ヌクレオチドを有し、配列表への記載を要する。

回答 : はい - PNA のプローブ (下の鎖)

この 2 本の鎖は完全には相補的ではないため、下の鎖も、個別の配列番号を割り当てて配列表に記載しなければならない。PNA すなわち「ペプチド核酸」を構成する個々の残基は、ST. 26 第 3 項(g)によればヌクレオチドとみなされる。したがって、下の鎖は、10 個以上の、数え上げられた「具体的に定義された」ヌクレオチドを有しており、配列表への記載を要する。

質問 3 : この配列を配列表にどの様に記載すべきか?

上の鎖は、以下のように配列表に記載しなければならない。

tagttcattgactaaggctccccattgactaaggcgactagcattgactaaggcaagc (配列番号 36)

下の鎖はペプチド核酸であるから、3 末端および 5 末端が存在しない。第 11 項によると、「5 末端から 3 末端の方向を模した左から右への方向で」配列表に記載しなければならない。したがって、本配列は以下の通り配列表に記載しなければならない。

cgcctnagtcaatggg (配列番号 37)

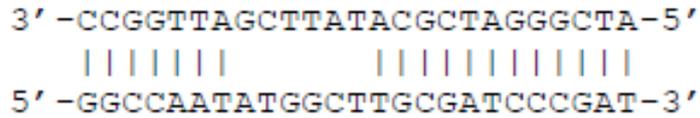
feature key「source」の qualifier「organism」には「synthetic construct」(合成コンストラクト)という値を、必須 qualifier「mol_type」には「other DNA」(他の DNA)という値を記載しなければならない。また、下の鎖は、feature key「modified_base」および必須 qualifier「mod_base」に「OTHER」の略語を用いて、feature table に記載しなければならない。qualifier「note」には「N-(2-aminoethyl) glycine nucleosides」(N-(2-アミノエチル)グリシンヌクレオシド)など、修飾塩基の省略されていない完全名称を記載しなければならない。

「n」残基は、feature key「modified_base」および必須 qualifier「mod_base」を略語「OTHER」と共に用いて、feature table に詳細を記載しなければならない。修飾塩基の省略されていない完全名称として、「N-(2-aminoethyl) glycine 5-nitroindole or N-(2-aminoethyl) glycine 3-nitroindole」(N-(2-アミノエチル)グリシン 5 ニトロインドールまたは N-(2-アミノエチル)グリシン 3 ニトロインドール)という qualifier「note」を記載しなければならない。

関連する ST. 26 の項 : 第 3 項(g)、第 7 項(a)、第 11 項(b)、第 17 項および第 18 項

例 11(b) -2: 二本鎖の塩基配列 - 塩基対のセグメントを持たないもの

特許出願には、以下の二本鎖の DNA 配列が記載されている。



質問 1： ST. 26 では上記配列の記載を要するか？

回答： はい

数え上げられた二本鎖の塩基配列の各鎖は、10 個以上の具体的に定義されたヌクレオチドを有している。二本鎖は互いに完全には相補的でないため、両方の鎖を、それぞれ個別の配列番号を割り当てて配列表に記載しなければならない。

質問 3： この配列を配列表にどの様に記載すべきか？

各鎖の配列は、5 末端から 3 末端の方向に表示し、それぞれ独自の配列番号を割り当てなければならない。

atcgggatcgcatattcgattggcc (top strand) (配列番号 38)

および、

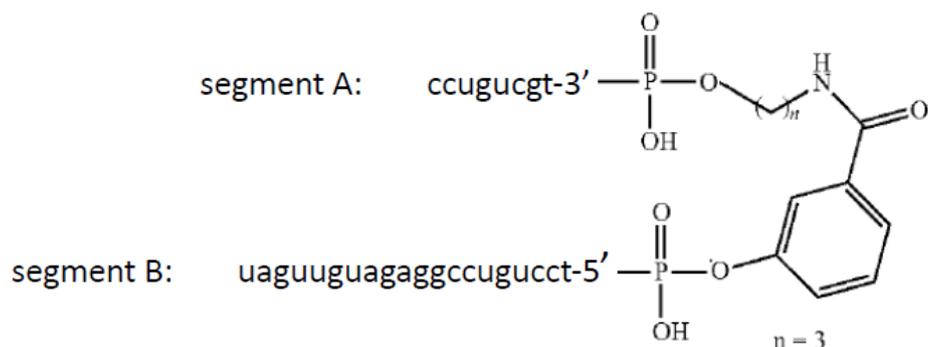
ggccaatatggcttgatcccgat (bottom strand) (配列番号 39)

関連する ST. 26 の項： 第 7 項(a)、第 11 項(b)および第 13 項

第 14 項 - RNA 中のウラシルを示すと解釈される記号「t」

例 14-1: RNA 中のウラシルを示す記号「t」

特許出願には、以下の化合物が記載されている。

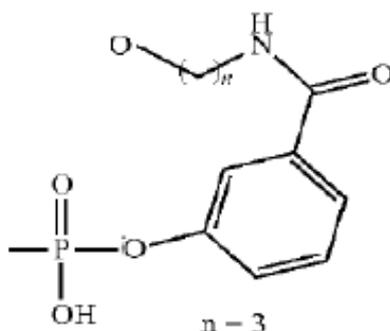


ここで、セグメント A およびセグメント B は RNA の配列である。

質問 1: ST. 26 では上記配列の記載を要するか?

回答: はい - セグメント B
いいえ - セグメント A

数え上げられた配列は、以下の「リンカー」構造によって分離された、具体的に定義されたヌクレオチドの 2 つのセグメントを有している。



リンカーの構造は、第 3 項(g)によれば、ヌクレオチドではないので、各セグメントは個別の配列であると考えなければならない。セグメント B は、10 個以上の具体的に定義されたヌクレオチドを含み、ST. 26 の第 7 項(a)では、配列表への記載を要する。セグメント A には 8 個の具体的に定義されたヌクレオチドしか含まれていないため、配列表への記載を要しない。

質問 2: ST. 26 に準拠する場合、本配列の配列表への記載は許容されるか?

回答: いいえ

セグメント A は、具体的に定義されたヌクレオチドが 10 個未満であるため、ST. 26 第 8 項のとおり、配列表に記載してはならない。

質問 3: この配列を配列表にどの様に記載すべきか?

セグメント B は RNA 分子であるため、「INSDSeq_moltype」の要素は「RNA」でなければならない。配列表において、RNA 分子のウラシルを示すために記号「u」を使用してはならない。第 14 項によると、記号「t」は RNA 中のウラシルと解釈される。したがって、セグメント B は以下の通りに配列表に記載しなければならない。

tcctgtccggagatgttgat (配列番号 40)

RNA 中のチミンは、修飾塩基、すなわち修飾を受けたウラシルとみなされ、配列内には「t」と示し、feature table に詳細に記載しなければならない。したがって、1 位のチミンは、feature key「modified_base」および値を「OTHER」として qualifier

「mod_base」、値を「thymine」として qualifier「note」を用いて更に記載しなければならない。可能であれば、qualifier「note」の値として、他の配列を直接示しても良い。

関連する ST. 26 の項：第 3 項(g)、第 7 項(a)、第 8 項、第 13 項、**第 14 項**、第 19 項および第 54 項

第 27 項 – アミノ酸を示す最も限定的な曖昧さ記号を用いるべき場合

例 27-1: アミノ酸配列の簡略式

(GGG_z)₂

上記の式の z は、任意のアミノ酸である。

質問 1: ST. 26 では上記配列の記載を要するか?

回答: はい

配列を数式として示している。(GGG_z)₂ は、GGGzGGGz という配列を簡略化して表現したものである。通常、まず配列を展開し、その後、任意の変数、すなわち「z」の定義を決定していた。

配列では、非準拠型の記号「z」を使用している。「z」の定義は、この記号を任意のアミノ酸と定義している本記述の配列の説明から判断しなければならない(本文書の「はじめに」参照)。この例では、「z」に対する制約、例えば、出現するたびに同じであることなどは示されていない。

この例のペプチドは、8 つの数え上げられたアミノ酸を有し、そのうちの 6 つは具体的に定義されたグリシン残基であり、残りの 2 つは、この配列では準拠型の記号「X」を使って表現されるべき変数「z」である。に定義された ST. 26 第 7 項(b)では、1 つの配列番号を持つ単一配列として、配列表へその配列を記載することを要する。

数え上げられた残基および具体的に定義された残基が連続していないにもかかわらず、この配列は第 7 項(b)を満たしているという点に留意されたい。

質問 3: この配列を配列表にどの様に記載すべきか?

この配列では、非準拠型の記号「z」が使用されており、明細書によれば、これは任意のアミノ酸である。「任意のアミノ酸」を示すために使用される準拠型の記号は「X」である。したがって、この配列は、以下の通り、単一の拡張した配列として示さなければならない。

GGGXGGGX (配列番号 41)

第 27 項によれば、「X」は、feature table に詳細が記載されている場合を除き、「A」、「R」、「N」、「D」、「C」、「Q」、「E」、「G」、「H」、「I」、「L」、「K」、「M」、「F」、「P」、「O」、「S」、「U」、「T」、「W」、「Y」または「V」のいずれか 1 つと解釈される。この例のおよび「X」は「任意のアミノ酸」を表しているので、feature key「VARIANT」および「X can be any amino acid」(X は任意のアミノ酸であり得る) という値の note qualifier を用いて注釈を付さなければならない。

可能であれば、それぞれの「X」は個別に注釈を付すべきである。しかし、連続した「X」残基の領域や、配列内に分散した多数の「X」残基は、存在位置記述子として構文「x...y」(x と y は先頭と末尾の「X」残基) を使用し、feature key「VARIANT」および値を「X can be any amino acid」(X は任意のアミノ酸であり得る) として qualifier「note」を用いて詳細な情報をまとめて記載しても良い。

さらにこの事例では、「z」が拡張配列の両方の位置で同じアミノ酸であるとは説明されていない。しかし「z」が両方の位置で同じアミノ酸であることが説明されたい場合、4 位と 8 位の「X」が両方の位置で同じであれば、どのようなアミノ酸であっても良いことを示す feature key「VARIANT」および qualifier「note」を記載すべきである。

関連する ST. 26 の項: 第 3 項(c)、第 7 項(b)および第 27 項

例 27-2: 簡略式 - 4 つ未満の具体的に定義されたアミノ酸

配列式(Gly-Gly-Gly-z)_n のペプチド。

明細書にはさらに、「z」は「任意のアミノ酸」と記載されており、

- (i) 変数「n」は、任意の長さである、または
- (ii) 変数「n」は 2～100 で、好ましくは 3 である

質問 1： ST. 26 では上記配列の記載を要するか？

回答： いいえ

式の数え上げられたペプチドの開示された実施形態(i)および(ii)の両方を考慮すると、「n」は「任意の長さ」であり得ることが明らかである。したがって、「n」の最も包括的な実施形態は不確定である。「n」が不確定であるため、式のペプチドを確定した長さに延長することはできず、したがって、簡略な式を検討しなければならない。

未展開式（「n」= 1）の数え上げられたペプチドは、それぞれがグリシン(Gly)である 3 つの具体的に定義されたアミノ酸と、記号「z」を記載する。従来、「Z」は「グルタミンまたはグルタミン酸」を表す記号であったが、実施例では、「z」を「任意のアミノ酸」と定義している（本文書の「はじめに」を参照のこと）。ST.26 では、具体的に定義されていないアミノ酸は「X」で示される。この分析によれば、数え上げられたペプチド、すなわち GGGX は具体的に定義された 4 つのアミノ酸を含んでいない。したがって、ST. 26 第 7 項(b)では、「n」もいくつかの実体形態では特定の値として定義されているにもかかわらず、記載する必要はない。

質問 2： ST. 26 に準拠する場合、本配列の配列表への記載は許容されるか？

回答： はい

この例では、変数「n」の特定の数値、すなわち、下限値が 2、上限値が 100、正確な値 3 が示されている。少なくとも 4 つの具体的に定義されたアミノ酸を含む任意の配列は、配列表に記載しても良い。

質問 3： この配列を配列表にどの様に記載すべきか？

GGGX が 100 回繰り返された配列が好ましい(配列番号 42)。さらなる注釈は、GGGX が 98 回まで繰り返された配列が削除され得ることを示すべきである。本発明の主要な部分であるさらなる特定の実施形態を、記載するよう強く推奨する。

第 27 項によると、記号「X」は、feature table に詳細が記載されている場合を除き、「A」、「R」、「N」、「D」、「C」、「Q」、「E」、「G」、「H」、「I」、「L」、「K」、「M」、「F」、「P」、「O」、「S」、「U」、「T」、「W」、「Y」、「V」のいずれか 1 つであると解釈する。配列番号 42 のおよび「X」は「任意のアミノ酸」を表しているので、feature key「VARIANT」および「X can be any amino acid」(X は任意のアミノ酸であり得る) という値の note qualifier を用いて注釈を付さなければならない。

可能であれば、それぞれの「X」は個別に注釈を付すべきである。しかし、連続した「X」残基の領域や、配列内に分散した多数の「X」残基は、存在位置記述子として構造「x...y」(x と y は先頭と末尾の「X」残基) を用いて feature key「VARIANT」および値を「X can be any amino acid」(X は任意のアミノ酸であり得る) とした qualifier「note」を用いて詳細な情報をまとめて記載しても良い。

注意：上記配列の好ましい表示は、特許出願日に配列表を提供するためのものであり、提供された情報が IPO により前の出願に主題事項を追加するとみなされる可能性がある関係上、特許出願日後に提供される配列表に同表示を適用できない場合がある。

関連する ST. 26 の項： 第 3 項(c)、第 7 項(b)、第 26 項および第 27 項

例 27-3: 簡略式 – 4 つ以上の具体的に定義されたアミノ酸

配列式(Gly-Gly-Gly-z)_n のペプチド。

上記の式の記号「z」は任意のアミノ酸であり、変数「n」は 2~100 で、3 が望ましい。

質問 1: ST. 26 では上記配列の記載を要するか?

回答: はい

式の数え上げられたペプチドは、それぞれがグリシン(Gly)である 3 つの具体的に定義されたアミノ酸と、記号「z」を記載している。従来、「Z」は「グルタミンまたはグルタミン酸」を表す記号であったが、この事例の記載では「z」を「任意のアミノ酸」と定義している(本文書の「はじめに」を参照のこと)。ST.26 では、具体的に定義されていないアミノ酸は「X」で表される。この分析によると、数え上げられたリピートペプチドは、4 つの具体的に定義されたアミノ酸を含んでいない。しかし、明細書では、変数「n」の特定の数値、すなわち、下限値が 2 で上限値が 100 と規定されている。したがって、本事例では、配列表に記載するために ST. 26 で求めている配列 GGGzGGGz に少なくとも 6 つの具体的に定義されたアミノ酸を有するペプチドが開示されている。

質問 3: この配列を配列表にどの様に記載すべきか?

記号「z」は任意のアミノ酸であることを示しているため、4 位および 8 位に存在するアミノ酸を表す準拠型の記号は「X」である。

ST. 26 では、残基ごとに数え上げられた単一配列のみを配列表に記載することが求められている。したがって、GGGX の 2 回、3 回または 100 回繰り返された 1 つの配列を配列表に記載しなければならないが、GGGX が 100 回繰り返された最も包括的な配列が好ましい(配列番号 42)(本文書の「はじめに」を参照のこと)。後者の場合、さらに注釈を付すことにより、GGGX が 98 回まで繰り返された配列が削除され得ることを示すことができる。それぞれ 2 回および 3 回繰り返された GGGX を含む 2 つの追加配列(配列番号 44 から 45)を記載することが強く推奨される。

第 27 項によれば、記号「X」は、feature table に詳細が記載されている場合を除き、「A」、「R」、「N」、「D」、「C」、「Q」、「E」、「G」、「H」、「I」、「L」、「K」、「M」、「F」、「P」、「O」、「S」、「U」、「T」、「W」、「Y」または「V」のいずれか 1 つであると解釈する。この例のおよび「X」は「任意のアミノ酸」を表しているため、feature key「VARIANT」および「X can be any amino acid」(X は任意のアミノ酸であり得る)という値の note qualifier を用いて注釈を付さなければならない。

可能であれば、それぞれの「X」は個別に注釈を付すべきである。しかし、連続した「X」残基の領域や、配列内に分散した多数の「X」残基は、存在位置記述子として構文「x...y」(x と y は先頭と末尾の「X」残基)を用いて、feature key「VARIANT」および値を「X can be any amino acid」(X は任意のアミノ酸とであり得る)として qualifier「note」を用いてまとめて記載しても良い。

さらに、この例では、変数「z」が、延長した配列の 2 つの出現箇所のそれぞれで同じであることを開示していない。しかし、「z」がすべての位置で同じアミノ酸であると説明されている場合、feature key「VARIANT」および qualifier「note」は、すべての位置で同じであれば、「X」がすべての位置で任意のアミノ酸となり得ることを示すべきである。

注意: 上記配列の好ましい表示は、特許出願日に配列表を提供するためのものであり、提供された情報が IPO により前の出願に主題事項を追加するとみなされる可能性がある関係上、特許出願日後に提供される配列表に同表示を適用できない場合がある。

関連する ST. 26 の項: 第 3 項(c)、第 7 項(b)、第 26 項および第 27 項

第 28 項 – 内部終結部位記号によって区切られたアミノ酸配列

例 28-1: コードする塩基配列、およびコードされたアミノ酸配列

特許出願には、以下の配列が記載されている。

caattcaggg tggtgaat atg gcg ccc aat acg caa acc gcc tct ccc cgc
Met Ala Pro Asn Thr Gln Thr Ala Ser Pro Arg

gcg ttg gcc gat tca tta atg cag ctg gca cga cag gtt tcc cga ctg
Ala Leu Ala Asp Ser Leu Met Gln Leu Ala Arg Gln Val Ser Arg Leu

Protein A

gaa agc ggg cag tga atg acc atg att acg gat tca ctg gcc gtc gtt
Glu Ser Gly Gln Met Thr Met Ile Thr Asp Ser Leu Ala Val Val

tta caa cgt cgt gac tgg gaa aac cct ggc gtt acc caa ctt aat cgc
Leu Gln Arg Arg Asp Trp Glu Asn Pro Gly Val Thr Gln Leu Asn Arg

Protein B

ctt gca gca cat tgg tgt caa aaa taa taataaccgg atgtactatt
Leu Ala Ala His Trp Cys Gln Lys

tatccctg atg ctg cgt cgt cag gtg aat gaa gtc gct taa gcaatcaatg
Met Leu Arg Arg Gln Val Asn Glu Val Ala

Protein C

tccgatgccg cgcgacgctt atccgaccaa catatcataa

質問 1 : ST. 26 では上記配列の記載を要するか?

回答 : はい

出願では、3 つの異なるアミノ酸配列をコードする、終止コドンを含む塩基配列が示されている。

数え上げられた塩基配列は、具体的に定義された 10 個以上のヌクレオチドが含まれており、1 本の配列として配列表に記載しなければならない。

コードされたアミノ酸配列については、第 28 項で、例えば「空白」などの内部終結部位記号で区切られたアミノ酸配列は、別の配列として記載しなければならないと定めている。「蛋白質 A」、「蛋白質 B」および「蛋白質 C」のそれぞれが、4 つ以上の具体的に定義されたアミノ酸を含んでいるため、ST. 26 第 7 項(b)では、それぞれ個別の配列番号を割り当てて、配列表に記載しなければならない

質問 3 : この配列を配列表にどの様に記載すべきか?

本塩基配列は、配列表に以下の配列として記載しなければならない。

caattcagggtggtgaatatggcgccaatacgcaaaccgcctctccccgcgctgtggccgattcattaatggaagcgggcagtgaatgaccatgattacg
gattcactggccgtcgttttacaacgtcgtgactgggaaaaccctggcgttacccaactaatcgccttgacgacattgggtgcaaaaataataataaccggat
gtactattatccctgatgctgcgtcaggtgaatgaagtcgcttaagcaatcaatgtcggatgcggcgcgacgcttatccgaccaacatataataa (配列
番号 46)

さらに、3つの蛋白質それぞれについて、feature key「CDS」を用いて核酸配列を記載すべきであり、「INSDFeature_location」要素で、終止コドンを含むそれぞれのコード領域の位置を特定しなければならない。また各 feature key「CDS」については、translation qualifier に、蛋白質のアミノ酸配列をその値として記載すべきである。本出特許では、翻訳に適用される遺伝コード表が開示されていない（附属書1、第9章、表7を参照のこと）。標準コード表が適用される場合は、qualifier「transl_table」は不要であるが、別の遺伝コード表が適用される場合は、qualifier「transl_table」に表7の適切な qualifier の値を示さなければならない。最後に、qualifier「protein_id」には、翻訳された各アミノ酸配列の配列番号を示す qualifier の値を記載しなければならない。

アミノ酸配列は、それぞれに配列番号を割り当てて、別の配列として記載しなければならない。

MAPNTQTASPRALADSLMQLARQVSRLESGQ (配列番号 47)

MTMITDSLAVVLQRRDWENPGVTQLNRLAAHWCQK (配列番号 48)

MLRRQVNEVA (配列番号 49)

注意：1つの配列として表される翻訳されたアミノ酸配列の説明については、「例 90-1 イントロンを有するコーディング配列によってコードされるアミノ酸配列」を参照のこと。

関連する ST. 26 の項：第7項、第26項、第28項、第57項、および第89項から第92項

第 29 項 - 「その他」のアミノ酸の表示

例 29-1: 「その他」のアミノ酸を示す最も限定的な曖昧さ記号

特許出願には、以下の配列が記載されている。

Ala-Hse-X1-X2-X3-X4-Tyr-Leu-Gly-Ser

上記の式で X1 はアラニン(Ala)またはグリシン(Gly)、
X2 はアラニン(Ala)またはグリシン(Gly)、
X3 はアラニン(Ala)またはグリシン(Gly)、
X4 はアラニン(Ala)またはグリシン(Gly)、
Hse はホモセリンである。

質問 1: ST. 26 では上記配列の記載を要するか?

回答: はい

数え上げられたペプチドは、具体的に定義された 5 つのアミノ酸を含んでいる。記号「X」は、選択的に 2 つのアミノ酸を表すために従来から使用されている（本文書の「はじめに」を参照のこと）。

具体的に定義された 5 つのアミノ酸、すなわち、アラニン(Ala)、チロシン(Tyr)、ロイシン(Leu)、グリシン(Gly)およびセリン(Ser)含まれるため、ST. 26 第 7 項(b)では、その配列を配列表に記載しなければならない。

質問 3: この配列を配列表にどの様に記載すべきか?

第 29 項は、「その他」のアミノ酸は記号「X」で表さなければならないとしている。本事例では、配列の 2 位にアミノ酸であるホモセリン(Hse)が存在しているが、これは附属書 1、第 3 節、表 3 にはない。したがって、ホモセリン(Hse)は「その他」のアミノ酸であり、記号「X」で表さなければならない。

X₁ から X₄ は、それぞれ A または G の可能性がある変異体の位置であり、アラニン(A)またはグリシン(G)の選択に対する最も限定的な曖昧さ記号は「X」である。したがって、この配列は次のように表しても良い。

AXXXXXYLGS (配列番号 50)

本文書の「はじめに」で述べたように、この発明の明細書または請求項に不可欠な特定の配列も記載することが強く推奨される。

アミノ酸であるホモセリン(Hse)は、附属書 1、第 4 節、表 4 にはないため、ST. 26 第 30 項に従い、feature key「SITE」および qualifier「note」を、ホモセリン(Hse)の省略されていない完全名称と共に記載しなければならない。

第 27 項によると、X₁ から X₄ は、2 つのアミノ酸の選択肢を示しているため、詳細に記載する必要がある。第 96 項によれば、feature key「VARIANT」は、qualifier「note」とその値「A or G」を用いて使用すべきである。ST. 26 第 34 項によれば、これらの位置は隣接しており、同じ記述であるため、要素「INSDFeature_location」の存在位置記述子として、構文「3..6」を用いて、まとめて注釈を付してもよい。

関連する ST. 26 の項: 第 3 項(a)、第 7 項(b)、第 25 項～第 27 項、第 29 項、第 30 項、第 34 項、第 66 項、第 70 項、第 71 項および、第 96 項から第 97 項

例 29-2: 対応する非修飾アミノ酸の使用

特許出願には、以下の配列が記載されている。

Ala-Hyl-Tyr-Leu-Gly-Ser-Nle-Val-Ser-5ALA

ここで、Hyl=ヒドロキシリジン（リジンの翻訳後修飾）、Nle=ノルロイシン、5ALA=δ-アミノレブリン酸とする。

質問 1：ST. 26 では上記配列の記載を要するか？

回答： はい

数え上げられたペプチドは、具体的に定義された 4 つ以上のアミノ酸を含んでいる。したがって、その配列は配列表に記載されていなければならない

質問 3：この配列を配列表にどの様に記載すべきか？

2 位のヒドロキシリジン、7 位のノルロイシン、10 位のδ-アミノレブリン酸はすべて「修飾アミノ酸」である。まず、それぞれの修飾アミノ酸について検討し、配列中に、対応する非修飾アミノ酸で表すべきか、変数「X」で表すべきかを判断しなければならない。第 29 項では、修飾アミノ酸は「可能な限り、対応する非修飾アミノ酸として配列中で表現されるべきである」としている。

修飾アミノ酸を、対応する非修飾の残基で表すか、変数「X」で表すかは、出願人の裁量に委ねられている。ただし、以下のガイドランスを考慮に入れるべきである。アミノ酸がメチル化やアセチル化などの部位の付加によって修飾されており、対応する非修飾アミノ酸の基本構造が概ね変化していない場合は、非修飾アミノ酸で表現することが推奨される。修飾アミノ酸が、対応する非修飾アミノ酸と構造的に大きく異なる場合は、「X」で表すことが推奨される。

ヒドロキシリジンの構造は、R 基の 3 番目の炭素がヒドロキシル基で修飾されていることを除けば、リジンとほぼ同じである。対応する非修飾リジン残基の基本構造はそのままのため、ヒドロキシリジンは配列上、「X」ではなくリジン（「K」）で表されるべきである。

ノルロイシンはロイシンの異性体である。ロイシンの R 基は炭素数 4 の鎖で、2 番目の炭素で分岐している。ノルロイシンも炭素数 4 の R 基を持っているが、直鎖状で、分岐していない。したがって、ノルロイシンは単にロイシンに修飾を加えただけのものではなく、（関連性はあるものの）全く異なる構造を持っている。そのため、配列表ではノルロイシンを「X」で表すことが推奨されている。

δ-アミノレブリン酸は、附属書 1 の表 3 に記載されているいずれのアミノ酸とも構造的に類似していない。したがって、δ-アミノレブリン酸は、配列表において「X」で表すことが推奨される。

したがって、この配列は、以下のように配列表に記載されるべきである。

AKYLG SXV SX（配列番号 51）

第 30 項では、各修飾アミノ酸のさらなる注釈を必要としている。

ヒドロキシリジンは、リジンの翻訳後修飾である。そのため、feature key「MOD RES」と qualifier「note」を用いて、修飾の内容を記述する必要がある。なお、「ヒドロキシリジン」は、附属書 1、第 4 節、表 4、「修飾アミノ酸のリスト」に記載されている。したがって、qualifier「note」の値には、省略されていない完全な名称「ヒドロキシリジン」ではなく、「Hyl」という略語を含めることができる。

ノルロイシンは翻訳後修飾残基ではないため、修飾を示す qualifier「note」とともに feature key「SITE」を用いて記述する必要がある。なお、「ノルロイシン」は附属書 1、第 4 節、表 4 にも記載されている。そのため、qualifier「note」の値には、省略されていない完全な名称「ノルロイシン」ではなく、「Nle」という略語を含めることができる。

δ-アミノレブリン酸もまた、翻訳後修飾残基ではないので、修飾を表す qualifier「note」とともに、feature key「SITE」を用いて記述しなければならない。δ-アミノレブリン酸は附属書 1 の第 4 節の表 4 には記載されていないため、qualifier「note」の値には、修飾残基の完全な省略されていない名称「δ-アミノレブリン酸」が含まれていなければならない。

関連する ST. 26 の項： 第 3 項(a)、第 3 項(e)、第 7 項(b)、第 29 項および、第 30 項

第 30 項 - 修飾アミノ酸の注釈

例 30-1: feature key「CARBOHYD」

特許出願には、グリコシル化された側鎖を含む、具体的に修飾されたアミノ酸を有するポリペプチドであって、ポリペプチドの 4 位および 15 位に存在するシステイン(Cys)のジスルフィド結合を形成していることを特徴とするポリペプチドが、以下の配列のように記載されている。

Leu-Glu-Tyr-Cys-Leu-Lys-Arg-Trp-Asn(シアリルオリゴ糖)-Glu-Thr-Ile-Ser-His-Cys-Ala-Trp

質問 1：ST. 26 では上記配列の記載を要するか？

回答： はい

数え上げられたペプチドは、具体的に定義された 17 個のアミノ酸を記載している。天然に存在するアミノ酸は 16 個であり、9 番目（アスパラギン）はグリコシル化されている。したがって、St. 26 第 7 項(b)で求められているように、この配列を配列表に記載しなければならない。

質問 3：この配列を配列表にどの様に記載すべきか？

ST. 26 第 29 項によると、修飾アミノ酸は可能な限り対応する非修飾アミノ酸として配列上に示すべきである。

したがって、配列は以下のように配列表に記載しなければならない。

LEYCLKRWNETHSHCAW (配列番号 52)

修飾アミノ酸は詳細な記載を要する。9 位のアスパラギンに糖鎖（アシアリルオリゴ糖）が結合している点を示すには、feature key「CARBOHYD」および qualifier「note」を使用すべきである。qualifier「note」は結合のタイプ(例 N-末端で結合)を値として記載する。要素「Feature location」の存在位置記述子には、修飾を受けたアスパラギンの残基位置番号を記載する。

また、2 つのシステイン(Cys)残基の間にはジスルフィド結合が存在する。したがって、feature key「DISULFID」を用いて鎖内クロスリンクを表現すべきである。feature location 要素は、リンクしているシステイン(Cys)残基の残基番号を「x..y」形式で表したもので、例えば「4..15」のようになる。必須の qualifier は、鎖内ジスルフィド結合を記述すべきである。

関連する ST. 26 の項： 第 3 項(a)、第 7 項(b)、第 26 項、第 29 項、第 30 項、第 66 項(c)、第 70 項および附属書 I、第 7 節および feature key7.4

第 36 項 – 正確な数の隣接する「n」または「X」残基の領域を含む配列

例 36-1: 単一配列として示される既知数の「X」残基の領域を 1 つ有する配列

LL-100-KYMR

アミノ酸のロイシンとリシンの間の「-100-」は、配列内の 100 のアミノ酸の領域を示している。

質問 1: ST. 26 では上記配列の記載を要するか?

回答: はい

ST. 26 第 36 項では、定義された数の「X」残基の 1 つ以上の領域で区切られた、少なくとも 4 つの具体的に定義されたアミノ酸を含む配列を記載することを要する。

この開示された配列では、非準拠型の記号、すなわち「-100-」を使用している。「-100-」の定義は、この記号をロイシンとリシンの間にある 100 個のアミノ酸と定義しており、開示された配列の説明から判断しなければならない（本文書の「はじめに」を参照のこと）。したがって「-100-」は、「X」残基の規定領域である。この配列の 106 個のアミノ酸のうち 6 つは具体的に定義されたアミノ酸であるため、ST. 26 第 7 項(b)では、この配列を配列表に記載しなければならない。

質問 3: この配列を配列表にどの様に記載すべきか?

非準拠型の記号「-100-」は、100 個の「X」残基を示している（アミノ酸を表すために使用される記号は、1 つの残基のみに相当するため）。したがって、LL と KYMR の間に 100 個の「X」残基を含む長さ 106 アミノ酸の 1 つの配列が、配列表に記載されなければならない（配列番号: 53）。

この配列は、LL と KYMR の間に 100 個の変数「X」を含んでいる。ST.26 のデフォルト値である「X」は、「A」、「R」、「N」、「D」、「C」、「Q」、「E」、「G」、「H」、「I」、「L」、「K」、「M」、「F」、「P」、「O」、「S」、「U」、「T」、「W」、「Y」、「V」のうちのいずれか 1 つである（第 27 項）。これら 100 個の変数「X」がこのデフォルト値以外で定義されている場合は、変数「X」それぞれに適切な注釈を付さなければならない。

関連する ST. 26 の項: 第 7 項(b)、第 26 項、第 27 項および第 36 項

例 36-2: 単一配列として示される既知の数の「X」残基の領域を複数有する配列

Lys-z₂-Lys-z_m-Lys-z₃-Lys-z_n-Lys-z₂-Lys

上記の式中の「z」は任意のアミノ酸であり、「m」=20、「n」=19 から 20、Z₂ はリシンのペアが任意の 2 つのアミノ酸で分離されていることを意味し、Z₃ はリシンのペアが任意の 3 つのアミノ酸で分離されていることを意味する。

質問 1: ST. 26 では上記配列の記載を要するか?

回答: はい

開示されている配列では、非準拠型の記号、すなわち「z」が使用されている。「z」は任意のアミノ酸と定義されている（本文書の「はじめに」を参照のこと）。任意のアミノ酸を表すために使用される準拠型の記号は「X」である。変数「X」の存在を考慮すると、ペプチドは、数え上げられ、具体的に定義された 6 つの残基(リシン)を含み、この配列は配列表への記載を要する。

質問 3: この配列を配列表にどの様に記載すべきか?

この配列では、非準拠型の記号「z」が使用されており、その定義は、この開示から判断しなければならない。「z」は任意のアミノ酸であると定義されているため、準拠型の記号は「X」である。好ましい、最も包含的な記載方法は以下の通りである（本文書の「はじめに」を参照のこと）。

KXXKXXXXXXXXXXXXXXXXXXXXXXXXKXXKXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXKXXK (配列番号 54)

ここで、z_n は、20 個の「X」に等しく、30 位に存在する変数「X」が欠失している可能性があることをさらに記述している。

代わりに、または上記に加えて、配列は以下の通りしても良い。

KXXKXXXXXXXXXXXXXXXXXXXXXXXXKXXKXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXKXXK (配列番号 55)

ここで、z_n は、19 個の「X」に等しく、29 位と 30 位の間に存在する変数「X」を挿入することができるというさらなる記述がある。

第 27 項によれば、記号「X」は、feature table に更に記述がある場合を除き、「A」、「R」、「N」、「D」、「C」、「Q」、「E」、「G」、「H」、「I」、「L」、「K」、「M」、「F」、「P」、「O」、「S」、「U」、「T」、「W」、「Y」または「V」のいずれか 1 つであると解釈する。配列番号 54 および 55 のおおよび「X」は「任意のアミノ酸」を表しているので、feature key「VARIANT」および「X can be any amino acid」(X は任意のアミノ酸であり得る) という値の note qualifier を用いて注釈を付さなければならない。

可能であれば、各「X」は、個別に注釈を付すべきである。しかし、連続した「X」残基の領域や、配列内に分散した多数の「X」残基は、存在位置記述子として構文「x...y」(「x」と「y」は先頭と末端の「X」残基の位置) を用いて、feature key「VARIANT」および qualifier「note」に値を「X can be any amino acid」(X は任意のアミノ酸であり得る) にまとめて記述しても良い。

関連する ST. 26 の項: 第 26 項、第 27 項および第 36 項

Gly-Gly----Gly-Gly-Xaa-Xaa

上記の式の記号----は配列内の未定義の間隙(ギャップ)であり、記号「Xaa」は任意のアミノ酸であり、グリシン残基および Xaa 残基はペプチド結合によって互いに結合している。

質問 1： ST. 26 では上記配列の記載を要するか？

回答： いいえ

ST. 26 第 37 項では、未定義の間隙(ギャップ)が存在するいかなる配列も記載してはならないことを求めている。

ST. 26 第 37 項では、未定義の間隙(ギャップ)に隣接する配列のうち、4 つ以上の具体的に定義されたアミノ酸を含む領域を記載することを求めている。上述の例では、未定義の間隙(ギャップ)に隣接するいずれかの領域は、各領域が 2 つの具体的に定義されたアミノ酸のみを含むため、記載を要しない。

質問 2： ST. 26 に準拠する場合、本配列の配列表への記載は許容されるか？

回答： いいえ - 配列の全体ではない。
 いいえ - 配列のどの領域でもない。

ST. 26 第 37 項は、配列全体を記載することを認めていない。

ST. 26 第 8 項では、未定義の間隙(ギャップ)に隣接するいずれの領域も記載することを認めていないが、これは各領域が 2 つの具体的に定義されたアミノ酸のみを含むためである。

関連する ST. 26 の項： 第 7 項(b)、第 8 項、第 26 項および第 37 項

例 37-2: 単一配列として示してはならない、未知の数の「X」残基の領域を有する配列

Gly-Gly----Gly-Gly-Ala-Gly-Xaa-Xaa

上記の式の記号----は配列内の未定義の間隙(ギャップ)であり、記号「Xaa」は任意のアミノ酸であり、グリシン残基および Xaa

残基はペプチド結合によって互いに結合している。

質問 1: ST. 26 では上記配列の記載を要するか?

回答: いいえ - 配列の全体ではない。
はい - 配列上の一領域

ST. 26 第 37 項では、未定義の間隙(ギャップ)を含むいかなる配列も記載することを認めていないが、しかし、未定義の間隙(ギャップ)に隣接する配列のうち、4 つ以上の具体的に定義されたアミノ酸を含むあらゆる領域を記載することを要する。

上記の例では、ST. 26 は、未定義の間隙(ギャップ)を含む配列全体と、未定義の間隙(ギャップ)に隣接するグリシン(Gly)-グリシン(Gly)領域 (2 つの具体的に定義されたアミノ酸のみを含む) の両方を記載することを要しない (そして認めていない)。しかし、ST.26 では、未定義の間隙(ギャップ)に隣接するグリシン(Gly)-グリシン(Gly)-アラニン(Ala)-グリシン(Gly)-Xaa-Xaa 領域は、少なくとも 4 つの具体的に定義されたアミノ酸を含んでいるため、記載することを要する。

質問 2: ST. 26 に準拠する場合、本配列の配列表への記載は許容されるか?

回答: いいえ - 配列全体も、Gly-Gly 領域も許容されていない

質問 3: この配列を配列表にどの様に記載すべきか?

未定義の間隙(ギャップ)に隣接する配列のうち、4 つの具体的に定義されたアミノ酸を含む配列は、以下のように表さなければならない。

GGAGXX (配列番号 58)

feature key「SITE」、feature location「1」および、例えば値を「This residue is linked N-terminally to a peptide having an N-terminal Gly-Gly and a gap of undefined length」(この残基は、グリシン-グリシンを含み、N 末端を有するペプチドおよび、未定義な長さの間隙(ギャップ)と N 末端で結合している) として「qualifier「note」を用いて、表された配列が、未定義の間隙(ギャップ)を含む、より長い配列の一部であることを示すよう注釈を付すべきである。

第 27 項によれば、記号「X」は、feature table にさらに記述される場合を除き、「A」、「R」、「N」、「D」、「C」、「Q」、「E」、「G」、「H」、「I」、「L」、「K」、「M」、「F」、「P」、「O」、「S」、「U」、「T」、「W」、「Y」または「V」のいずれか 1 つであると解釈する。配列番号 58 のおよび「X」は「任意のアミノ酸」を表しているので、feature key「VARIANT」および「X can be any amino acid」(X は任意のアミノ酸であり得る) という値の note qualifier を用いて注釈を付さなければならない。

可能であれば、各「X」は個別に注釈を付すべきである。しかし、連続した「X」残基の領域や、配列内に分散した多数の「X」残基は、存在位置記述子として構文「x...y」(x と y は先頭と末尾の「X」残基の位置) として、feature key「VARIANT」および値を「X can be any amino acid」(X は任意のアミノ酸であり得る) として qualifier「note」を用いてまとめて記載しても良い。

関連する ST. 26 の項: 第 7 項(b)、第 8 項、第 26 項、第 27 項および第 37 項

第 55 項 – DNA と RNA の両方を含む塩基配列

例 55-1: DNA/RNA 分子の結合

特許出願には、以下のオリゴヌクレオチドの配列が記載されている。：

AGACCTTcggagucuccuguugaacagauagucaaaguagauC

上記の式で、大文字で示されているのは DNA の残基であり、小文字で示されているのは RNA の残基である。

質問 1：ST. 26 では上記配列の記載を要するか？

回答：はい

開示されている配列は、10 個以上の数え上げられた具体的に定義されたヌクレオチドを有しているため、配列表への記載を要する。

質問 3：この配列を配列表にどの様に記載すべきか？

この塩基配列は、以下の通り配列表に記載しなければならない。

agaccttcggagtctcctgttgaacagatagtagtcaaagtagatc (配列番号 93)

ウラシルヌクレオチドは、配列表に記号「t」で表さなければならない点に注意する。

ST. 26 第 55 項では、DNA セグメントおよび RNA セグメントの両方を含む塩基配列は、分子タイプを「DNA」としなればならず、feature key「source」および必須 qualifier「organism」の値を「synthetic construct」および必須 qualifier「mol_type」の値を「other DNA」（他の DNA）として、さらに記述しなければならない。また、配列の各セグメントは、セグメントの位置を含む feature key「misc_feature」および、そのセグメントが DNA であるのか RNA であるかを示す qualifier「note」を用いてさらに記述しなければならない。この開示された配列には、2 つの DNA セグメント(ヌクレオチドの位置は 1 位から 7 位および 43 位)、および 1 つの RNA セグメント(ヌクレオチドの位置は 8 位から 42 位)が含まれている。

関連する ST. 26 の項：第 7 項、第 14 項、第 55 項～第 56 項および第 83 項

第 89 項 – feature key「CDS」

例 89-1: コードする塩基配列、およびコードされたアミノ酸配列

特許出願には、以下のような塩基配列とその翻訳が記載されている。

atg acc gga aat aaa cct gaa acc gat gtt tac gaa att tta tga

Met Thr Gly Asn Lys Pro Glu Thr Asp Val Tyr Glu Ile Leu STOP

質問 1 : ST. 26 では上記配列の記載を要するか?

回答 : はい

数え上げられた塩基配列は、10 個以上の具体的に定義されたヌクレオチドを有している。

数え上げられたアミノ酸配列は、4 個以上の具体的に定義されたアミノ酸を有している。

質問 3 : この配列を配列表にどの様に記載すべきか?

本塩基配列は、以下の通り配列表に記載しなければならない。

atgaccggaataaacctgaaaccgatgtttacgaaattttatga (配列番号 59)

この塩基配列は、さらに feature key「CDS」を用いて記載すべきであり、要素 INSDFeature_location は、終止コドンを含む配列の全体（すなわち 1 位から 45 位まで）を特定しなければならない。また、qualifier「translation」の値は「MTGNKPETDVYEIL」であるべきである。この出願では、翻訳に適用される遺伝コード表が開示されていない(附属書 1、第 9 節、表 7 を参照のこと)。標準コード表が適用される場合は、qualifier「transl_table」は必要ないが、別の遺伝コード表が適用される場合は、qualifier「transl_table」は、表 7 で示されている適切な値と共に用いなければならない。最後に、qualifier「protein_id」には、値として翻訳後のアミノ酸配列の配列番号を使用しなければならない。

アミノ酸配列には、以下のような一文字のコードを用いて、個別に配列番号を表示しなければならない。

MTGNKPETDVYEIL (配列番号 60)

数え上げられたアミノ酸配列の直後の終止コドンは、配列表のアミノ酸配列に記載してはならない。

注意 : 上記配列の好ましい表示は、特許出願日に配列表を提供するためのものであり、提供された情報が IPO により前の出願に主題事項を追加するとみなされる可能性がある関係上、特許出願日後に提供される配列表に同表示を適用できない場合がある。

関連する ST. 26 の項 : 第 7 項(a)、第 7 項(b)、第 26 項、第 28 項、第 89 項、第 90 項および第 92 項

例 89-2: 開示された配列を超える feature location

特許出願には、部分的なコード配列およびその翻訳後アミノ酸配列を示す以下の図が含まれている。

```

cat cac gca gca gaa tgt gga ttt tgt cct caa caa tgg caa gtt cta      48
His His Ala Ala Glu Cys Gly Phe Cys Pro Gln Gln Trp Gln Val Leu
1                               5                               10                               15

cgt ggg agt ctg tgc att tgt gag ggt cca gct gaa gga tgg ttc ata      96
Arg Gly Ser Leu Cys Ile Cys Glu Gly Pro Ala Glu Gly Trp Phe Ile
                               20                               25                               30

tca aga tgt tgg tta tgg tgt ggg cct caa gtc caa ggc ttt atc ttt      144
Ser Arg Cys Trp Leu Trp Cys Gly Pro Gln Val Gln Gly Phe Ile Phe
                               35                               40                               45

gga gaa ggc aag gaa gga ggc ggt gac aga cgg gct gaa gcg agc cct      192
Gly Glu Gly Lys Glu Gly Gly Gly Asp Arg Arg Ala Glu Ala Ser Pro
                               50                               55                               60

cag gag ttt tgg gaa tgc act tgg                                     216
Gln Glu Phe Trp Glu Cys Thr Trp
65                               70

```

図 1—442 個のアミノ酸で構成され、ITCH1 蛋白質の 20 位から 91 位のアミノ酸をコードしている、ホモ・サピエンスの ITCH1 遺伝子の部分的なコード配列

質問 1: ST. 26 では上記配列の記載を要するか?

回答: はい

本出願では、塩基配列およびその翻訳後のアミノ酸配列を開示している。数え上げられた塩基配列は、具体的に定義された 10 個以上のヌクレオチドを含み、配列表に記載しなければならない。

アミノ酸配列は、具体的に定義された 4 つ以上のアミノ酸を有しており、個別の配列番号を持つ別の配列として、配列表に記載しなければならない。

質問 3: この配列を配列表にどの様に記載すべきか?

この塩基配列は、以下のように配列表に記載しなければならない。

```

catcacgcagcagaatgtggattttgtcctcaacaatggcaagttctacgtgggagctgtgcatgtgagggtccagctgaaggatggtcatatcaagatgttg
ggttatggtgtggcctcaagtccaaggctttatctttggagaaggcaaggaaggaggcggtgacagacgggctgaagcgagccctcaggagtttgggaatg
cacttgg (配列番号 94)

```

この塩基配列は、さらに feature key「CDS」を用いて記載すべきである。また、要素「INSDFeature_location」は、配列内の feature の location を特定しなければならず、終止コドンを含まなければならない。

上掲の図では、開始コドンまたは終止コドンを含まない、部分的なコード配列を示している。しかし、配列の説明によると、開始コドンは 1 位のヌクレオチドの上流に存在し、終止コドンは 216 位の最後のヌクレオチドの下流に存在することを示している。

ST. 26 では、要素「INSDSeq_sequence」内の配列の範囲を超える残基のナンバリングを存在位置記述子に記載してはならないと規定している。したがって、上記の例では、feature key「CDS」の存在位置記述子には、1 位から 216 位の範囲外の位置番号を記載することはできない。要素「INSDSeq_sequence」内の終止コドンの位置は、終止コドンが 216 位の下流に存在することを示すために、記号「>」を用いて示さなければならない。同様に、開始コドンの位置が 1 位の上流に存在することを示すために、記号「<」を用いて表すことができる。したがって、feature key「CDS」の存在位置記述子は以下のように表されるべきである。

<1..>216

なお、「<」および「>」は予約文字であり、配列表の XML インスタンスでは、それぞれ「<」および「>」に置き換えられる。qualifier「translation」の値は、蛋白質のアミノ酸配列を値として記載すべきである。

この図では、翻訳に適用される遺伝コード表がが開示されていない(附属書 1、第 9 節、表 7)。標準コード表が適用される場合には、qualifier「transl_table」は不要であるが、別の遺伝コード表が適用される場合には、ST.26 附属書 1 の表 7 から適切な値を qualifier「transl_table」に示さなければならない。最後に、feature key「CDS」の qualifier「protein_id」には、qualifier 値「翻訳後のアミノ酸配列の配列番号」を用いなければならない

翻訳後のアミノ酸配列は、個別の配列番号を持つ別の配列として記載しなければならない。

HHAECGFCPQQWQVLRGSLCICEGPAEGWFISRCWLWCGPQVQGFIFGEGKEGGDRRAEASPQEFWECT
W (配列番号配列番号 95)

注意：上記配列の好ましい表示は、特許出願日に配列表を提供するためのものであり、提供された情報が IPO により前の出願に主題事項を追加するとみなされる可能性がある関係上、特許出願日後に提供される配列表に同表示を適用できない場合がある。

関連する ST. 26 の項：第 7 項、第 41 項、第 65 項、第 66 項、第 70 項、第 71 項、第 89 項および第 92 項

第 92 項 - コード配列によりコードされたアミノ酸配列

例 92-1: イントロンと結合したコード配列によりコードされたアミノ酸配列

特許出願には、コード配列およびその翻訳を示す次の図が記載されている。

```

atg aag act ttc gca gcc ttg ctt tcc gct gtc act ctc gcg ctc tcg
Met Lys Thr Phe Ala Ala Leu Leu Ser Ala Val Thr Leu Ala Leu Ser

gtg cgc gcc cag gcg gct gtc tgg agt caa t gtaagtgccg ctgcttttca
Val Arg Ala Gln Ala Ala Val Trp Ser Gln

ttgatacgag actctacgcc gagctgacgt gctaccgtat ag gt gcc ggt aca
Cys Gly Gly Thr

ccg ggt tgg acg gcc gag acc act tgc gtt gct ggt tcg gtt tgt acc
Pro Gly Trp Thr Gly Glu Thr Thr Cys Val Ala Gly Ser Val Cys Thr

tcc ttg agc tca gtgagcgact ttcaatccgt cgtcattgct cctcatgtat
Ser Leu Ser Ser

tgacgattgg ccttcatag tca tac tct caa tgc gtt ccg gcc tcc gca acg
Ser Tyr Ser Gln Cys Val Pro Gly Ser Ala Thr

tcc agc gct ccg gcg gcc ccc tca gcg aca act tca gcc ccc gca cct
Ser Ser Ala Pro Ala Ala Pro Ser Ala Thr Thr Ser Gly Pro Ala Pro

acg gac gga acg tgc tcg gcc agc ggg gca tgg ccg cca ttg acc tga
Thr Asp Gly Thr Cys Ser Ala Ser Gly Ala Trp Pro Pro Leu Thr Ter

```

図 1 - 太字で示されているヌクレオチドはイントロン領域

質問 1 : ST. 26 では上記配列の記載を要するか?

回答 : はい

この特許出願では、塩基配列およびそのアミノ酸翻訳を開示している。数え上げられた塩基配列は、具体的に定義された 10 個以上のヌクレオチドを有しているため、単一配列として記載しなければならない。

塩基配列は、非コード配列 (イントロン) で区切られたコード配列 (エキソン) を有している。図は、塩基配列の翻訳を、非連続的な 3 つのアミノ酸配列として表している。図の説明によると、太字のヌクレオチド領域は、蛋白質に翻訳される前に RNA 転写物からスプライシングされるイントロン配列である。したがって、3 つのアミノ酸配列は、実際には、4 つ以上の具体的に定義されたアミノ酸を有する単一の連続して数え上げられた配列であり、1 つの配列として記載しなければならない。

質問 3 : この配列を配列表にどの様に記載すべきか?

本塩基配列は、以下のように、配列表に記載しなければならない。

```

atgaagactttgcagccttgccttccgctgtcactctcgcgctctcgggtgcgcgccaggcggtgtctggagtaaatgtaagtgccgctgctttcattgatagca
gactctacgccgagctgacgtgtaccgtataggtggcggtacaccgggtggacgggagaccacttgcgttgcgtggttcggtttgtacctcctgagctcag
tgagcgactttcaatccgctgctcattgctcctcatgtattgacgattggcctcatagtcatactctcaatgcgttccgggctccgcaacgtccagcgctccggcgcc
cccctcagcgacaacttcaggccccgcacctacggacggaacgtgctcggccagcggggcatggccgcatgacctga (配列番号 75)

```

塩基配列は、feature key「CDS」を用いてさらに記述すべきである。また、要素「INSDFeature_location」は、「Ter」で示される終止コドンを含むコード配列の location を示さなければならない。CDS INSDFeature_location は、存在位置演算子「join」を

用いて、「join(x1...y1,x2...y2,x3...y3)」の形式、すなわち「join(1...79,142...212,272...400)」の形式で、示された位置でコードされた翻訳産物が結合し、単一の連続したポリペプチドを形成していることを示さなければならない。また、translation qualifierとしてその値に蛋白質のアミノ酸配列を指定すべきである。(なお、配列の末尾の位置にある終結部位記号「Ter」は、アミノ酸配列に含めてはならない)。この出願では、翻訳に適用される遺伝コード表を開示しない(附属書 1、第 9 節、表 7 参照)。「標準コード」表が適用される場合は、qualifier「tran_table」は不要だが、異なる遺伝コード表が適用される場合は、qualifier「tran_table」に表 7 からの適切な qualifier 値を示さなければならない。最後に、qualifier「protein_id」には、翻訳後アミノ酸配列の配列番号を示す qualifier 値を入力しなければならない。

アミノ酸配列は、単一配列として記載しなければならない。

MKTFAALLSAVTLALSVRAQA AVWSQCGGTPGWTGETTCVAGSVCTSLSSSYSQCVPGSATSSAPAAPSAT
TSGPAPTDGTCSASGAWPPLT (配列番号 76)

関連する ST. 26 の項：第 7 項、第 26 項、第 28 項、第 57 項、第 67 項および第 89 項から第 92 項

第 93 項 – それぞれ残基により数え上げられた、一次配列と変異体を含む配列

例 93-1: 数え上げられた変異体の表示

明細書には、以下の配列のアラインメントが記載されている。

```

D. melanogaster   ACATTGAATCTCATACCACTTT
D. virilis        ...-..G...C...-.G.....
D. simulans       GT..G.CG..GT..SGT.G...

```

質問 1: ST. 26 では上記配列の記載を要するか?

回答: はい

この技術分野では、「この位置に存在する残基は、その上に存在する残基と同じである」と示すために、配列アラインメントに「ドット」を記載する慣習が一般的である。したがって、クロショウジョウバエおよびオナジショウジョウバエの配列内の「ドット」は、ある位置が、キロショウジョウバエと同じヌクレオチドであることを示す単なる略式の方法であるため、数え上げられた且体的に定義されたヌクレオチドであることを示している。また、配列のアラインメントでは、アラインメントを最大化するために、残基が存在しないことを示す記号「-」が示されることが多い。

従って、キロショウジョウバエとオナジショウジョウバエの塩基配列は、22 個の数え上げられた、且体的に定義されたヌクレオチドを有するが、クロショウジョウバエの塩基配列は 19 個である。したがって、ST.26 第 7 項(a)では、それぞれの配列に別々の配列番号を割り当てて配列表に記載することを要する。

質問 3: この配列を配列表にどの様に記載すべきか?

キロショウジョウバエの配列は、以下のように配列表に記載しなければならない。

acattgaatctcataccacttt (配列番号 61)

クロショウジョウバエの配列は以下の通り配列表に記載しなければならない。

acatggatcccacgacttt (配列番号 62)

オナジショウジョウバエの配列は以下の通り配列表に記載しなければならない。

gtatggcgtcgtatsgtagttt (配列番号 63)

関連する ST. 26 の項: 第 7 項(a)、第 13 項および第 93 項

例 93 -2: 数え上げられた変異体の表示

明細書には、ペプチドおよびその機能的変異体が以下の表が含まれている。以下の表の空欄は、変異体中のアミノ酸が「配列」中の対応するアミノ酸と同じであることを示し、「-」は「配列」中の対応するアミノ酸の欠失を示している。

| 位置 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|----------|---|---|---|---|---|---|---|---|---|
| 配列 | A | V | L | T | Y | L | R | G | E |
| 機能的変異体 1 | | | | | | | | | A |
| 機能的変異体 2 | | | P | | | P | | | |
| 機能的変異体 3 | | | A | I | G | Y | | | |
| 機能的変異体 4 | | | | | | | - | | |

質問 1 : ST. 26 では上記配列の記載を要するか?

回答 : はい

上述の通り、この表の空欄は、変異体のアミノ酸が、「配列」の対応するアミノ酸と同じであることを示している。したがって、変異体の配列のアミノ酸は数え上げられ、具体的に定義されている。

4 つの変異体の配列は、それぞれ 4 つ以上の数え上げられた、具体的に定義されたアミノ酸が含まれているため、ST. 26 第 7 項 (b)により、別々の配列番号を割り当てたうえで、配列表に記載しなければならない。

示されるように、この表の空白は、変異体のアミノ酸が「配列」の対応するアミノ酸と同じであることを示す。したがって、変異体配列のアミノ酸は数え上げられ、具体的に定義されている。

4 つの変異体配列は、それぞれ 4 つ以上の数え上げられた、具体的に定義されたアミノ酸を含んでいるので、各配列は ST.26 第 7 項(b)により、別々の配列番号を割り当てて配列表に記載することを要する。

質問 3 : この配列を配列表にどの様に記載すべきか?

AVLTYLRGE (配列番号 77)

AVLTYLRGA (配列番号 78)

AVPTYPRGE (配列番号 79)

AVAIGYRGE (配列番号 80)

AVLTYLGE (配列番号 81)

関連する ST. 26 の項 : 第 7 項(b)、第 26 項および第 93 項

例 93 -3: 共通配列の表示

特許出願には、以下の複数の配列のアラインメントを示す図 1 が記載されている。

| | |
|-------------|-----------------------------------|
| 共通配列 | LEGnEQFINAakIIRHPkYnrkTlnNDImLIK |
| ホチサプトンス | LEGNEQFINAAKIIIRHPQYDRKTLNNDIMLIK |
| スマトラオランウータン | LEGNEQFINAAKIIIRHPQYDRKTVNNDIMLIK |
| アヌビスヒ | LEGTEQFINAAKIIIRHPDYDRKTLNNDILLIK |
| キンシコウ | LEGTEQFINAAKIIIRHPNYNRITLDNDILLIK |
| ボノボ | LEGNEQFINAAKIIIRHPKYNRITLNNDIMLIK |
| ウンナンシシバナザル | LEGNEQFINATKIIIRHPKYNGNTLNNDIMLIK |
| キンシコウ | LEGNEQFINATQIIIRHPKYNGNTLNNDIMLIK |

共通配列には、保存されたアミノ酸残基を表す大文字が含まれており、小文字の「n」、「a」、「k」、「r」、「l」、「m」は、アラインメントされた配列の中で優勢なアミノ酸残基を表している。

質問 1：ST. 26 では上記配列の記載を要するか？

回答：はい

共通配列の小文字は、それぞれ 1 つのアミノ酸残基を表している。その結果、共通配列は、図 1 の残りの 7 つの配列と同様に、少なくとも 4 つの具体的に定義されたアミノ酸が含んでいる。ST. 26 第 7 項(b)では、8 つの配列すべてを配列表に記載することを要する。

質問 3：この配列を配列表にどの様に記載すべきか？

共通配列の小文字は、特定の位置に存在する変異体の中で優性なアミノ酸を表現するための曖昧さ記号として使用されている。したがって、小文字の「n」、「a」、「k」、「r」、「l」、「m」は、準拠型の記号を非準拠型の方法で使用したものであり、共通配列は、それぞれの小文字の代わりに曖昧さ記号を使用して表さなければならない。

曖昧さ記号は、最も限定的なものを使用すべきである。共通配列のほとんどの位置では、「X」が最も限定的な曖昧さ記号である。しかし、20 位および 25 位の「D」または「N」に対する最も限定的な曖昧さの記号は「B」である。共通配列は、以下のように配列表に記載すべきである。

LEGXEQFINAXXIIIRHPXYBXXTXBNDIXLIK (配列番号 82)

第 27 項によれば、記号「X」は、feature table にさらに記述がある場合を除き、「A」、「R」、「N」、「D」、「C」、「Q」、「E」、「G」、「H」、「I」、「L」、「K」、「M」、「F」、「P」、「O」、「S」、「U」、「T」、「W」、「Y」、「V」のいずれか 1 つであると解釈する。したがって、共通配列の各「X」は、feature key「VARIANT」および各位置で考えられる変異体を示す qualifier「note」を用いて、feature table にさらに記述しなければならない。

LEGNEQFINAAKIIIRHPQYDRKTLNNDIMLIK (配列番号 83)

LEGNEQFINAAKIIIRHPQYDRKTVNNDIMLIK (配列番号 84)

LEGTEQFINAAKIIIRHPDYDRKTLNNDILLIK (配列番号 85)

LEGTEQFINAAKIIIRHPNYNRITLDNDILLIK (配列番号 86)

LEGNEQFINAAKIIIRHPKYNRITLNNDIMLIK (配列番号 87)

LEGNEQFINATKIIIRHPKYNGNTLNNDIMLIK (配列番号 88)

LEGNEQFINATQIIIRHPKYNGNTLNNDIMLIK (配列番号 89)

注意：上記配列の好ましい表示は、特許出願日に配列表を提供するためのものであり、提供された情報が IPO により前の出願に主題事項を追加するとみなされる可能性がある関係上、特許出願日後に提供される配列表に同表示を適用できない場合がある。

関連する ST. 26 の項：第 7 項(b)、第 26 項、第 27 項、第 93 項および第 97 項

第 94 項 – 数え上げられた代替残基を有する単一配列として開示される変異体配列

例 94 -1: 数え上げられた代替アミノ酸による単一配列の表示

特許出願では、以下の配列のペプチドを請求範囲としている。

(i) Gly-Gly-Gly-[Leu or Ile]-Ala-Thr-[Ser or Thr]

質問 1：ST. 26 では上記配列の記載を要するか？

回答：はい

この配列は、4 つの具体的に定義されたアミノ酸を有しており、ST. 26 第 7 項(b)では、配列表に配列を記載することを要する。

質問 3：この配列を配列表にどの様に記載すべきか？

附属書 1、第 3 節、表 3 では、曖昧さ記号「J」はイソロイシンまたはロイシンと定義している。したがって、この配列は以下のように配列表に記載するのが望ましい。

GGGJATX (配列番号 64)

「X」がセリンまたはトレオニンであることを示すために、feature key「VARIANT」および qualifier「note」を用いて、feature table に更に記述することを要する。

また、配列は、例えば以下のように記載しても良い。

GGGLATS (配列番号 65)

これは、feature key「VARIANT」および、L を I に、S を T に置き換えることができることを示すために qualifier「note」を用いて、feature table に記述することを要する。

注意：上記配列の好ましい表示は、特許出願日に配列表を提供するためのものであり、提供された情報が IPO により前の出願に主題事項を追加するとみなされる可能性がある関係上、特許出願日後に提供される配列表に同表示を適用できない場合がある。

関連する ST. 26 の項：第 7 項(b)、第 8 項、第 26 項、第 27 項、第 94 項および第 97 項

第 95 項(a) – 複数の独立した変異体を有する一次配列を参照することによってのみ開示される変異した配列

例 95(a) -1: 一次配列の注釈による変異体配列の表示

特許出願には、以下の配列が記載されている。

「Peptide fragment 1 is Gly-Leu-Pro-Xaa-Arg-Ile-Cys wherein Xaa can be any amino acid....

In another embodiment, peptide fragment 1 is Gly-Leu-Pro-Xaa-Arg-Ile-Cys wherein Xaa can be Val, Thr, or Asp....

In another embodiment, peptide fragment 1 is Gly-Leu-Pro-Xaa-Arg-Ile-Cys wherein Xaa can be Val.]

(ペプチドセグメント 1 は、Gly-Leu-Pro-Xaa-Arg-Ile-Cys であり、Xaa は任意のアミノ酸であり得る....

別の実施形態では、ペプチドセグメント 1 は、Gly-Leu-Pro-Xaa-Arg-Ile-Cys であり、ここ k で、Xaa は、Val、Thr または Asp....である。

別の実施例では、ペプチドセグメント 1 は、Gly-Leu-Pro-Xaa-Arg-Ile-Cys であり、ここで Xaa は Val であることができる。)

質問 1：ST. 26 では上記配列の記載を要するか？

回答：はい

開示された 3 つの実施形態のそれぞれにおける「ペプチドセグメント 1」は、少なくとも 6 つの具体的に定義されたアミノ酸を含んでいる。したがって、ST. 26 第 7 項(b)で求められているように、配列表に記載しなければならない。

質問 3：この配列を配列表にどの様に記載すべきか？

この例では、「ペプチドセグメント 1」の数え上げられた配列が、3 つの異なる実施形態として、それぞれ Xaa の選択的な記述を伴って 3 回開示されている。この例では、「X」が Xaa の位置を表す最も限定的な曖昧さ記号である。

ST. 26 は、開示された、数え上げられた配列を一度だけ記載することを求めている。3 つの実体形態のうち、最も包括的なものでは、Xaa は任意のアミノ酸である（本文書の「はじめに」を参照のこと）。したがって、配列表に記載しなければならない配列は、以下の通りである。：

GLPXRIC (配列番号 66)

第 27 項によれば、記号「X」は、feature table にさらに記述がある場合を除き、「A」、「R」、「N」、「D」、「C」、「Q」、「E」、「G」、「H」、「I」、「L」、「K」、「M」、「F」、「P」、「O」、「S」、「U」、「T」、「W」、「Y」または「V」のいずれか 1 つであると解釈する。配列番号 66 の「X」は「任意のアミノ酸」を表しているので、feature key「VARIANT」および「X can be any amino acid」(X は任意のアミノ酸であり得る) という値の note qualifier を用いて注釈を付さなければならない。

可能であれば、それぞれの「X」は個別に注釈を付すべきである。しかし、連続した「X」残基の領域や、配列内に分散した多数の「X」残基は、存在位置記述子として構文「x...y」(x と y は先頭と末尾の「X」残基の位置) を用いた feature key「VARIANT」および値を「X can be any amino」(X は任意のアミノ酸であり得る) とした qualifier「note」を用いてさらに記述しても良い。

本文書の「はじめに」で述べたように、この発明の開示または請求項に不可欠な追加の配列を記載することが強く推奨される。

上記の例では、以下の追加の 3 つの配列を、それぞれ個別の配列番号を割り当てて配列表に記載することが強く推奨される。

GLPVRIC (配列番号 67)

GLPTRIC (配列番号 68)

GLPDRIC (配列番号 69)

注意：上記配列の好ましい表示は、特許出願日に配列表を提供するためのものであり、提供された情報が IPO により前の出願に主題事項を追加するとみなされる可能性がある関係上、特許出願日後に提供される配列表に同表示を適用できない場合がある。

関連する ST. 26 の項：第 7 項(b)、第 26 項、第 27 項および第 95 項(a)

第 95 項(b) – 複数の相互に依存する変異体を有する一次配列を参照することによってのみ開示される変異体配列

例 95(b) -1: 複数の相互に依存する変異体を有する、個々の変異体配列の表示

特許出願には、以下の共通配列が記載されている。

cgaaatgn₁cccactacgaatgn₂cacgaatgn₃cccaca

上記の式の、n₁、n₂および n₃ は、a、t、g または c であり得る。

幾つかの変異した配列が以下のように開示されている。

n₁ が a であれば、n₂ および n₃ は t、g または c である。

n₁ が t であれば、n₂ および n₃ は a、g または c である。

n₁ が g であれば、n₂ および n₃ は t、a または c である。

n₁ が c であれば、n₂ および n₃ は t、g または a である。

質問 1：ST. 26 では上記配列の記載を要するか？

回答：はい

本配列は、10 個以上の数え上げられた「具体的に定義された」ヌクレオチドを有しているため、ST. 26 第 7 項(a)により、配列表への記載を要する。

質問 3：この配列を配列表にどの様に記載すべきか？

数え上げられた配列は、具体的に定義された 10 個以上のヌクレオチドおよび 3 つの「n」残基が存在する。ST. 26 は、開示された、数え上げられた配列を記載することを求めており、曖昧さ記号が適切な場合には、最も限定的な記号を使用すべきである。この例では、n₁、n₂、n₃ は、a、t、g、c のいずれかであるため、「n」が最も限定的な曖昧さ記号となる。したがって、本配列は以下の通り配列表に記載しなければならない。

cgaaatgncccactacgaatgncacgaatgncccaca (配列番号 70)

ST.26 の第 15 項は、記号「n」は、feature table に詳細な記述がある場合を除き、「a」、「c」、「g」、「t/u」のいずれか 1 つと解釈されるとする。この配列のすべての残基の値「n」は、デフォルトの「a」、「c」、「g」、または「t」と同等であるため、さらなる注釈は不要である。

数え上げられた配列は、3 つの異なる位置に変異体を有し、変異の発生は相互に依存している。本文書の「はじめに」で述べたとおり、この発明の重要な部分である追加の実施形態を表す追加の配列を記載することが強く推奨される。したがって、ST. 26 第 95 項(b)によれば、追加の実施形態は、それぞれが個別の配列番号を割り当てた 4 つの個別の配列として配列表に記載すべきである。

cgaaatgaccactacgaatgbcacgaatgbcaccaca (配列番号 71)

cgaaatgtcccactacgaatgvcacgaatgvccaccaca (配列番号 72)

cgaaatggcccactacgaatghcacgaatghccaccaca (配列番号 73)

cgaaatgccccactacgaatgdcacgaatgdccaccaca (配列番号 74)

(b は t、g または c を示し、v は a、g または c を示し、h は t、a または c を示し、d は t、g または a を示している点に注意する (附属書 I、第 1 節、表 1 を参照のこと)。

ST. 26 第 15 項によると、変数の位置を表すためには、最も限定的な記号を使用しなければならない。したがって、n₂ および n₃ は、配列内の「n」で表してはならない。

注意：上記配列の好ましい表示は、特許出願日に配列表を提供するためのものであり、提供された情報が IPO により前の出願に主題事項を追加するとみなされる可能性がある関係上、特許出願日後に提供される配列表に同表示を適用できない場合がある。

関連する ST. 26 の項：第 7 項(a)、第 15 項および第 95 項(b)

[次頁から付録が続く]

付 録

ガイダンス文書、XML 形式での配列

本ガイダンス文書は、以下のリンク先で参照可能。

https://www.wipo.int/standards/en/xml_material/st26/st26-annex-vi-appendix-guidance-document-sequences.xml

[次頁から附属書 VII が続く]

附属書 VII

WIPO 標準 ST. 25 から ST. 26 への配列表の移行に関する推奨事項：
(移行の際に) 追加または削除される可能性のある主題事項

Version 1.5

2021 年 11 月 5 日開催の第 9 回 WIPO 標準委員会(CWS) 会合で採択

はじめに

塩基配列およびアミノ酸配列の表示に関する要件は、WIPO 標準 ST. 25 と ST. 26 とで異なる。そのため、ST.26 では、優先権主張出願に裏付けされていない可能性のある、ST.26 に基づく国際出願の一部として提出された配列表において、主題事項の追加や削除が必要になるのではないかという疑問が生じている。

本標準の範囲(scope)

本文書は、ST. 26 の必須要件、およびこれらの要件に起因する潜在的影響について述べている。ST. 25 の配列表に記載される情報の ST.26 における表示方法が明らかではない場合は、主題事項が削除されないように、常に出願明細書にその情報を記載することができる。

追加または削除される可能性のある主題事項に関する推奨

本文書に記載された事項を検討した結果、ST. 25 から ST. 26 への移行だけでは、特に ST.25 の配列表が標準 ST.25 に完全に準拠している場合には、本質的に主題事項の追加や削除にはならないことがわかっている。しかし、出願人の注意を要するであろうシナリオは存在するため、主題事項の追加や削除を避けるための推奨事項を示すこととする。

シナリオ 1

ST.25 は、出願人名を表す <110> のように、様々なタイプのデータのタグ付けに数値による識別子を使用している。ST.26 では、データのタグ付けに、英語の用語を要素名や属性として使用している。

推奨

ST. 26 で用いられている用語は、単にデータ内容の種類を表すものであり、ST. 26 の要素名や属性を使用しても、主題事項が追加されたことにはならない。

シナリオ 2

ST. 26 では、(a) 分岐配列、(b) D-アミノ酸を含む配列、(c) ヌクレオチド類似体および、(d) 脱塩基部位を有する配列も、配列表内に記載するよう明示している。ST. 25 では、このような配列の記載、あるいは禁止の要件は明確ではない。

推奨

ST. 25 の配列表内に含まれていなかった場合でも、ST.26 の配列表にこれらの配列を表示するには、出願時の内示内容で十分である。ST. 26 が必要とする特定の種類の情報に関しては、開示されている以上の主題事項を追加しないように注意しなければならない、例. 塩基配列に用いる qualifier「mol_type」に関しては後述(シナリオ 4)を参照のこと。

シナリオ 3

ST. 26 では、10 個未満の具体的に定義されたヌクレオチド(「n」を含まず)、4 個未満の具体的に定義されたアミノ酸(「X」を含まず)の配列は除かれている。

推奨

除かれた配列は、その配列がまだ出願の主要部に記載されていない場合、そこへの記載もできる。

シナリオ 4

ST. 26 では、すべての塩基配列およびすべてのアミノ酸配列に「source」という必須の feature key があり、それぞれに 2 つの必須 qualifier を有している。ST.25 では、対応する塩基配列の feature key はある (ほとんど使用されていない) が、その qualifier はなく、対応するアミノ酸配列の feature key はない。

塩基配列

ST. 26 – feature key 5.37 「source」およびその必須 qualifier 6.45 「organism」および 6.39 「mol_type」 (ST.26 の第 75 項を参照のこと)。

| qualifier | qualifier 値 |
|-----------|--------------------------------|
| mol_type | genomic DNA |
| | genomic RNA |
| | mRNA |
| | tRNA |
| | rRNA |
| | other DNA (合成分子に適用) |
| | other RNA (合成分子に適用) |
| | transcribed RNA |
| | viral cRNA |
| | unassigned DNA (生体分子が不明な場合に適用) |
| | unassigned RNA (生体分子が不明な場合に適用) |

アミノ酸配列

ST. 26 – feature key 7.30 「source」および、その必須 qualifier 8.3 「organism」および 8.1 「mol_type」 (ST. 26 の第 75 項を参照のこと)。

| qualifier | qualifier 値 |
|-----------|-------------|
| mol_type | protein |

推奨

唯一懸念される問題は、塩基配列に関して用いる qualifier「mol_type」値の統制語彙である。上掲の qualifier 値の一部は、本開示で十分に説明されていない可能性がある。配列に最も一般的な qualifier 値を用いれば、主題事項の追加にはならないこともある (例. 合成分子の場合は「other DNA」および「other RNA」を選択し、生体分子の場合は「unassigned DNA」および「unassigned RNA」を選択する)。

シナリオ 5

ST.25 では、配列に「Xaa」が含まれる場合、<221> feature name 欄と<222> feature location 欄に付随する<223>欄に、その残基に関する詳細な情報の記載を要する。ST.25 では、「Xaa」 (ST.26 では「X」) のデフォルト値を規定していないが、ST.26 ではそのようなデフォルト値が用意されているため、詳細情報は必ずしも必要ではない。ペプチド配列で最も頻繁に使用される 2 つの注釈は、変数「Xaa」または「X」に関する「any amino acid」または「any natural occurring amino acid」である。この表現は、ST.25 または ST.26 に示されているアミノ酸の表に記載されていないアミノ酸を含むと解釈されかねない。つまり ST.26 のデフォルト値である「X」は、附属書 1 に記載されている 22 種類の個々のアミノ酸のいずれかである (第 3 章、表 3 参照)。この ST.26 のデフォルト値は、それ自身が追加または削除された主題事項を構成する可能性があり、従って、ST.25 から ST.26 への移行時に特許出願の範囲に悪影響を及ぼす可能性がある。

推奨

- (a) ST. 25 の配列表に、「Xaa」に対応する<221> feature name、<222> feature location および「Xaa」に関する<223> further information が含まれており、<221> feature name が ST. 26 の適切な feature key (例. 「SITE」、 「VARIANT」または「UNSURE」) でもある場合、ST.26 の feature key を使用すべきである。また移行により主題事項が削除されないように、<223>欄の情報を、付随する qualifier「note」に記載しなければならない。
- (b) ST. 25 の配列表に、「Xaa」に対応する<221> feature name、<222> feature location および「Xaa」に関する<223> further information に関する詳細な情報が記載されており、<221> feature name が ST. 26 の feature key では無い場合、必要に応じて ST.26 の feature key「SITE」または「REGION」を使用すべきである。さらに、移行により主題事項が削除されないように、不適切な<221> feature name ならびに<223>欄の情報を、付随する qualifier「note」に記載しなければならない。例えば ST. 25 の配列表では、ST. 25 や ST. 26 にない feature name である<221>Variable が<223>が任意のアミノ酸であるとの詳細な情報とともに使用される。この例で ST. 26 の qualifier「note」の値は、「Variable – Xaa is any amino acid」となる。
- (c) ST.25 の配列表に、Xaa に対応する<221>、<222>または<223>欄がない場合、または Xaa に対応する<221>欄および<222>欄はあるものの、対応する<223>欄に情報がない場合 (いずれのシナリオも ST.25.に準拠していないが、それにもかかわらず発生している)、「Xaa」を記載するために出願書類に含まれる情報は、ST.26 の qualifier「note」に、SITE、REGION、UNSURE などの適切な feature key や location とともに記載すべきである。

シナリオ 6

ST. 25 では、配列内でウラシルは記号「u」で表し、チミンは記号「t」で表している。ST. 26 では、ウラシルとチミンは両方とも記号

「t」で表し、さらなる注釈を加えない。すなわち「t」は、RNA 中のウラシルと DNA 中のチミンの両方を表す。

推奨

- (a) DNA の配列にはウラシルが含まれている場合、ST. 26 ではそれを修飾塩基とみなし、記号「t」で示し、feature key「modified_base」およびその qualifier「mod_base」(qualifier 値は「OTHER」)と qualifier「note」(qualifier の値は「uracil」)を用いて記載するよう求めている。この ST. 26 の注釈は、ST.25 の DNA 配列に「u」が記載されている場合には、追加された主題事項とはみなされない。
- (b) RNA の配列がチミンを含む場合、ST. 26 はそれを修飾塩基とみなし、記号「t」で示し、feature key「modified_base」およびその qualifier「mod_base」(qualifier 値は「OTHER」)と qualifier「note」(qualifier 値は「thymine」)を用いて記載することを必要としている。この ST.26 の注釈は、ST. 25 の RNA 配列に「t」が記載されている場合には、追加された主題事項とはみなされない。

シナリオ 7

ST. 25 および ST. 26 では共に、修飾塩基または修飾を受けたアミノ酸には詳細を記述しなければならない。ST. 26 では、修飾塩基の同一性を、該当する場合は附属書 1、第 2 節、表 2 の略語を用いて示しても良いとしている。それ以外の場合は、修飾塩基の省略されていない完全名称を表示しなければならない。同様に修飾を受けたアミノ酸の同一性は、該当する場合は、附属書 1、第 4 節、表 4 の略語を用いて示しても良い。それ以外の場合は、修飾を受けたアミノ酸の省略されていない完全名称を表示しなければならない。対照的に、修飾を受けた残基が ST. 25 の表に記載されていない場合は、完全な省略されていない名称を使用する必要はなく、代わりに略語が使用されることも少なくない。

推奨

- (a) 修飾塩基、または修飾を受けたアミノ酸に関して、附属書 1、第 2 節、表 2 または第 4 節、表 4 に記載されていない略語のみが、出願および ST. 25 の配列表の両方で使用されており、その略語が 1 つの特定の修飾塩基または修飾を受けたアミノ酸のみを示すものであるとして当業者に知られている場合、省略されていない完全名称を使用すること自体は、追加された主題事項とはならない。
- (b) 特許出願書類および ST. 25 に準拠型の配列表の両方に、ST. 26 の附属書 1、第 2 節、表 2 または第 4 節、表 4 には記載されていない修飾を受けたヌクレオチドまたは修飾を受けたアミノ酸のいずれか一方の略称が用いられ（そして、特許出願書類には化学構造が記載されていない場合）、そしてその略称は該当する技術分野で 1 つの特定の修飾を受けたヌクレオチドまたは修飾を受けたアミノ酸を示しているとは知られていない場合（すなわち、その略称が該当する技術分野で全く知られていないか、または複数の異なる修飾を受けたヌクレオチドまたは修飾を受けたアミノ酸を表すものを思われるかのいずれか一方）、この状況では主題事項を追加せずに ST. 26 への準拠を確保することは不可能である。この場合はもちろん、優先権を有する特許出願書類および配列表そのものが不明瞭である。主題事項が削除されるという可能性を排除するためには、ST. 25 に準拠する配列表に記載されているその略称を、ST. 26 の qualifier「note」の値として用いると共に、その修飾を受けたヌクレオチドまたは修飾を受けたアミノ酸の完全で、省略されていない名称も qualifier 値として用いるべきである。但し、ST. 26 で必要としている修飾塩基または修飾を受けたアミノ酸の完全で、省略されていない名称を用いて ST. 26 に準拠型の配列表を作成したとしても、ST. 26 に準拠する配列表には先に提出された特許出願書類が有する優先権は発生しない。したがって将来問題にならないように、ST. 25 に準拠する特許出願書類の配列表および明細書を作成する際は省略されていない名称を記載するよう注意すべきである。

シナリオ 8

ST. 25 には、ST. 26 では記載されていない複数の feature key が記載されている。したがって出願人は、ST.25 の feature key に記載されている情報を、主題事項を追加または削除することなく、ST. 26 に準拠型の方法で取り込むように注意しなければならない。

推奨

以下の表は、旧 ST.25 の feature key に記載されている情報を、主題事項が追加または削除されることなく ST. 26 に準拠型の配列表に記載できる方法についてのガイダンスを示している。番号 1 から 23 は塩基配列に用いる feature key であり、番号 24 から 43 はアミノ酸配列に用いる feature key である。

| No. | ST.25 Feature key <221> | ST.26 equivalent | | |
|-----|---|-----------------------------|-------------------------------|---|
| | | Feature key | Qualifier | Qualifier value |
| 1 | allele | misc_feature | allele | <223> value |
| 2 | attenuator | regulatory ¹ | regulatory_class ¹ | “attenuator” |
| | | | note (if <223> present) | <223> value |
| 3 | CAAT_signal | regulatory ¹ | regulatory_class ¹ | “CAAT_signal” |
| | | | note (if <223> present) | <223> value |
| 4 | conflict | misc_feature | note | “conflict” and <223> value |
| 5 | enhancer | regulatory ¹ | regulatory_class ¹ | “enhancer” |
| | | | note (if <223> present) | <223> value |
| 6 | GC_signal | regulatory ¹ | regulatory_class ¹ | “GC_signal” |
| | | | note (if <223> present) | <223> value |
| 7 | LTR | mobile_element ¹ | rpt_type ¹ | “long_terminal_repeat” |
| | | | note (if <223> present) | <223> value |
| 8 | misc_signal | regulatory ¹ | regulatory_class ¹ | “other” |
| | | | note (if <223> present) | <223> value |
| 9 | mutation | variation | note | “mutation” and <223> value |
| 10 | old_sequence | misc_feature | note | “old_sequence” and <223> value |
| 11 | polyA_signal | regulatory ¹ | regulatory_class ¹ | “polyA_signal_sequence” |
| | | | note (if <223> present) | <223> value |
| 12 | promoter | regulatory ¹ | regulatory_class ¹ | “promoter” |
| | | | note (if <223> present) | <223> value |
| 13 | RBS | regulatory ¹ | regulatory_class ¹ | “ribosome_binding_site” |
| | | | note (if <223> present) | <223> value |
| 14 | repeat_unit (a) when repeat_region not used | misc_feature | note | “repeat_unit” and <223> value |
| | repeat_unit (b) when repeat_region used | repeat_region | rpt_unit_range | 1 st residue..last residue |
| 15 | satellite | repeat_region | satellite | “satellite” (or “microsatellite” or “minisatellite” – if supported) |
| | | | note (if <223> present) | <223> value |
| 16 | scRNA | ncRNA ¹ | ncRNA_class ¹ | “scRNA” |
| | | | note (if <223> present) | <223> value |
| 17 | snRNA | ncRNA ¹ | ncRNA_class ¹ | “snRNA” |
| | | | note (if <223> present) | <223> value |
| 18 | TATA_signal | regulatory ¹ | regulatory_class ¹ | “TATA_box” |
| | | | note (if <223> present) | <223> value |
| 19 | terminator | regulatory ¹ | regulatory_class ¹ | “terminator” |
| | | | note (if <223> present) | <223> value |
| 20 | 3'clip | misc_feature | note | “3'clip” and <223> value |
| 21 | 5'clip | misc_feature | note | “5'clip” and <223> value |
| 22 | -10_signal | regulatory ¹ | regulatory_class ¹ | “minus_10_signal” |
| | | | note (if <223> present) | <223> value |
| 23 | -35_signal | regulatory ¹ | regulatory_class ¹ | “minus_35_signal” |
| | | | note (if <223> present) | <223> value |

注釈 1 --- ST.26 では、特定の ST.25 の feature (例: TATA_signal) を、より広義の feature key/qualifier/value (例: regulatory/regulatory_class/TATA_box) に置き換えることを求める場合がある。このような場合、狭義の ST.25 の feature は、先願が優先されるが、ST.26 の feature key/qualifier (例 regulatory/regulatory_class) がより広義である場合は、先願が優先されることはない。

| No. | ST. 25 featurekey<221> | ST. 26 で ST. 25 feature key<221>と同等のもの | | |
|-----|------------------------|---|-----------|---|
| | | Feature key | qualifier | qualifier 値 |
| 24 | NON_CONS | この feature は 1 本の配列に含まれる未知数の残基で構成される間隙(ギャップ)に関する feature であり、このような間隙 (ギャップ) が配列上に 1 つあるいはそれ以上の数存在し配列を区切っている場合は、1 本の配列として配列表に記載することは ST. 25 (第 22 項) でも、ST. 26 (第 37 項) でも禁止されている。したがって、間隙 (ギャップ) で区切られた各領域が具体的に定義された残基で構成された領域で、ST. 26 の第 7 項の定めに該当する場合は、配列表に個別の配列として含めなければならない、それぞれ個別の配列番号を割り振らなければならない。また、主題事項が追加/削除されないように、その様な各配列はより長い配列の一部であり、長さが定義されていない間隙 (ギャップ) を含んでいることを示す注釈も付さなければならない。 | | |
| | | REGION | note | Description |
| | | 配列がどの位置で何と結合しているかに関する記載、例. 本残基は、グリシン-グリシンを含み N-末端を有するペプチドおよび、長さが定義されていない間隙(ギャップ)と N-末端で結合している。 | | |
| 25 | SIMILAR | REGION | note | 「SIMILAR」および、存在する場合は <223>の値 |
| 26 | THIOETH | CROSSLINK | note | 「THIOETH」および、存在する場合は <223>の値 |
| | | 詳しい location に関するガイダンスについては、ST. 26 の附属書 1、第 7 節、アミノ酸配列の feature key「CROSSLINK」(7.9)およびその Comment を参照のこと。 | | |
| 27 | THIOLEST | CROSSLINK | note | 「THIOLEST」および、存在する場合は <233>の値 |
| | | 詳しい location に関するガイダンスについては、ST. 26 の附属書 1、第 7 節、アミノ酸配列の feature key「CROSSLINK」(7.9)およびその Comment を参照のこと。 | | |
| 28 | VARSP LIC | 以降のシナリオ 13 で検討する。 | | |
| 29 | ACETYLATION | MOD_RES | note | 「ACETYLATION」および、存在する場合は <223>の値 |
| | | | note | (主題事項の追加無しで)可能な場合は、ST. 26 の附属書 1、第 7 節、アミノ酸配列の feature key「MOD_RES」(7.18) comment で必要とされている情報を値とする。 |
| 30 | AMIDATION | MOD_RES | note | 「AMIDATION」および、存在する場合は <223>の値 |
| | | | note | (主題事項の追加無しで)可能な場合は、ST. 26 の附属書 1、第 7 節、アミノ酸配列の feature key「MOD_RES」(7.18) comment で必要とされている情報を値とする。 |
| 31 | BLOCKED | MOD_RES | note | 「BLOCKED」および、存在する場合は <223>の値 |
| | | | note | (主題事項の追加無しで)可能な場合は、ST. 26 の附属書 1、第 7 節、アミノ酸配列の featurekey「MOD_RES」(7.18)およびその Comment で必要とされている情報を値とする。 |
| 32 | FORMYLATION | MOD_RES | note | 「FORMYLATION」および、存在する場合は <223>の値 |
| | | | note | (主題事項の追加無しで)可能な場合は、ST. 26 の附属書 1、第 7 節、アミノ酸配列の feature key「MOD_RES」(7.18) comment で必要とされている情報を値とする。 |

| | | | | |
|----|--|---------|------|---|
| 33 | GAMMA-CARBOXYGLUTAMIC ACID HYDROXYLATION | MOD_RES | note | 「GAMMA-CARBOXYLGLUTAMIC ACID HYDROXYLATION」および、存在する場合は<223>の値 |
| | | | note | (主題事項の追加無しで)可能な場合は、ST. 26 の附属書 1、第 7 節、アミノ酸配列の feature key「MOD_RES」(7.18) comment で必要とされている情報を値とする。 |
| 34 | METHYLATION | MOD_RES | note | 「METHYLATION」および、存在する場合は<223>の値 |
| | | | note | (主題事項の追加無しで)可能な場合は、ST. 26 の附属書、第 7 節、アミノ酸配列の feature key「MOD_RES」(7.18) comment で必要とされている情報を値とする。 |
| 35 | PHOSPHORYLATION | MOD_RES | note | 「PHOSPHORYLATION」および、存在する場合は<223>の値 |
| | | | note | (主題事項の追加無しで)可能な場合は、ST. 26 の附属書 1、第 7 節、アミノ酸配列の feature key「MOD_RES」(7.18) comment で必要とされている情報を値とする。 |
| 36 | PYRROLIDONE CARBOXYLIC ACID | MOD_RES | note | 「PYRROLIDONE CARBOXYLIC ACID」および、存在する場合は<223>の値 |
| | | | note | (主題事項の追加無しで)可能な場合は、ST. 26 の附属書 1、第 7 節、アミノ酸配列の feature key「MOD_RES」(7.18) comment で必要とされている情報を値とする。 |

| No. | ST. 25
featurekey<221> | ST. 26 で ST. 25 featurekey<221>と同等のもの | | |
|-----|---------------------------|---------------------------------------|-----------|--|
| | | Featurekey | qualifier | qualifier 値 |
| 37 | SULFATATION | MOD_RES | note | 「SULFATATION」および、存在する場合は<223>の値 |
| | | | note | (主題事項の追加無しで)可能な場合は、ST. 26 の附属書 1、第 7 節、アミノ酸配列の feature key「MOD_RES」(7.18) comment で必要とされている情報を値とする。 |
| 38 | MYRISTATE | LIPID | note | 「MYRISTATE」および、存在する場合は<223>の値 |
| | | | note | (主題事項の追加無しで)可能な場合は、ST. 26 の附属書 1、第 7 節、アミノ酸配列の feature key「LIPID」(7.16) comment で必要とされている情報を値とする。 |
| 39 | PALMITATE | LIPID | note | 「PALMITATE」および、存在する場合は<223>の値 |
| | | | note | (主題事項の追加無しで)可能な場合は、ST. 26 の附属書 1、第 7 節、アミノ酸配列の feature key「LIPID」(7.16) comment で必要とされている情報を値とする。 |
| 40 | FARNESYL | LIPID | note | 「FARNESYL」および、存在する場合は<223>の値 |
| | | | note | (主題事項の追加無しで)可能な場合は、ST. 26 の附属書 1、第 7 節、アミノ酸配列の featurekey「LIPID」(7.16) およびその comment で必要とされている情報を値とする。 |
| 41 | GERANYL-GERANYL | LIPID | note | 「GERANYL-GERANYL」および、存在する場合は<223>の値 |
| | | | note | (主題事項の追加無しで)可能な場合は、ST. 26 の附属書 1、第 7 節、アミノ酸配列の feature key「LIPID」(7.16) comment で必要とされている情報を値とする。 |
| 42 | GPI-ANCHOR | LIPID | note | 「GPI-ANCHOR」および、存在する場合は<223>の値 |
| | | | note | (主題事項の追加無しで)可能な場合は、ST. 26 の附属書 1、第 7 節、アミノ酸配列の feature key「LIPID」(7.16) comment で必要とされている情報を値とする。 |
| 43 | N-ACYL
DIGLYCERIDE | LIPID | note | 「N-ACYL DIGLYCERIDE」および、存在する場合は<223>の値 |
| | | | note | (主題事項の追加無しで)可能な場合は、ST. 26 の附属書 1、第 7 節、アミノ酸配列の feature key「LIPID」(7.16) comment で必要とされている情報を値とする。 |

シナリオ 9

ST. 25 および ST. 26 の両方に存在する一部の feature key(塩基配列に用いる feature key およびアミノ酸配列に用いる feature key の両方)それぞれには、以下の必須 qualifier を用いるよう ST. 26 では求められている。例えば、塩基配列に用いる feature key「modified_base」は、ST. 25 および ST. 26 の両方に存在する。しかしシナリオ 7 に適切な推奨が記載されている。ST. 25 では qualifier は定められていないが、<223>のフリーテキストの欄がその機能を担っている。したがって、ST. 25 の<223>欄に記載されている情報が、ST. 26 の必須 qualifier の qualifier 値として適切である場合は、qualifier 値として用いるべきである。ST. 25 が定める<223>欄が記載されていない場合か、または ST. 26 の必須 qualifier の qualifier 値として適切ではない情報が記載されているかのいずれかである場合、出願人は、ST. 25 の feature key<223>欄に記載されている情報を、主

題事項を追加または削除することなく、ST. 26 に準拠型の方法で ST. 26 の必須 qualifier の qualifier 値として用いるよう注意しなければならない。

塩基配列²

| Featurekey | 必須 qualifier |
|---------------------|--------------------|
| 5.12 - misc_binding | 6.3 - bound_moiety |
| 5.30 - protein_bind | 6.3 - bound_moiety |

推奨

- (a) ST. 25 の feature key<233>欄の記載がないか記載された情報が不適切である場合、出願明細書に feature location が示す位置で核酸と結合する可能性のある分子/複合体の名称が明らかにされている場合、その名称を qualifier 「bound_moiety」に記載すべきである。
- (i) ST. 26 の qualifier「bound_moiety」への移行が不適切な情報が ST. 25 の feature key<223>の欄に記載されている場合は、適切で使用が任意の qualifier、例、「note」に記載すべきである。
- (b) ST. 25 の feature key<233>欄の記載がないか記載された情報が不適切である場合、出願明細書に feature location が示す位置で核酸と結合する可能性のある分子/複合体の名称が明らかにされていない場合、feature key 「misc_binding」または「protein_bind」の代わりに、ST.26 feature key「misc_feature」を用い、qualifier「note」を付すべきである。
- (i) ST. 25 の feature key<223>欄の記載がない場合、qualifier「note」の値は、ST.25 の feature key の名称を用いるべきである。
- (ii) ST. 25 の feature key<223>欄に不適切な情報が記載されている場合、qualifier「note」の値は ST. 25 の feature key の名称および、<223>欄に記載されている情報とすべきである。

アミノ酸配列³

| Feature key | 必須 qualifier |
|-----------------|--------------|
| 7.2 – BINDING | 8.2 – note |
| 7.4 – CARBOHYD | 8.2 – note |
| 7.10 – DISULFID | 8.2 – note |
| 7.11 – DNA_BIND | 8.2 – note |
| 7.12 – DOMAIN | 8.2 – note |
| 7.16 – LIPID | 8.2 – note |
| 7.17 – METAL | 8.2 – note |
| 7.18 - MOD_RES | 8.2 – note |
| 7.23 – NP_BIND | 8.2 – note |
| 7.29 – SITE | 8.2 – note |
| 7.39 – ZN_FING | 8.2 – note |

推奨

- (a) ST. 25 の feature key<233>欄の記載がないか記載された情報が不適切である場合、出願明細書に必須 qualifier に必要な特定の情報が明らかにされている場合、その情報は qualifier「note」に記載すべきである。
- (i) ST.25 <223>欄に記載された情報のうち、必須 qualifier「note」（Feature key の定義とコメントを参照）に含めるのに不適切なものは、2 つ目の qualifier「note」に挿入すべきである。

² 以下の表の数字は、ST.26、附属書 I 統制語彙の feature key および qualifier の番号を示している。

³ 以下の表の数字は、ST.26、附属書 I 統制語彙の feature key および qualifier の番号を示している。

- (b) ST. 25 の feature key<233>欄の記載がないか記載された情報が不適切である場合、出願明細書に必須 qualifier で必要とされる特定の情報が明らかにされていない場合、ST. 26 の feature key「SITE」(1つのアミノ酸について) または feature key「REGION」(アミノ酸の範囲について)を、qualifier「note」とともに代わりに用いるべきである。
- (i) ST. 25 の feature key<223>欄の記載がない場合、qualifier「note」の値は、ST.25 の feature key の名称 (すなわち<223>) とすべきである。
- (ii) ST. 25 の feature key<223>欄に不適切な情報が記載されている場合、qualifier「note」の値は、ST. 25 の feature key の名称と<223>欄の情報とすべきである。

シナリオ 10

ST. 25 の特定の feature key にはそれぞれ、feature location を示す<222> 欄が記載されているが、ST. 25 ではほとんど feature location の表示を必要とせず、location のフォーマットも標準化されていない。また ST. 25 には「join」等の存在位置演算子が存在しない。ST. 26 には標準化された存在位置記述子および存在位置演算子が存在し、feature それぞれに少なくとも1つの存在位置記述子を用いなければならない (feature「CDS」は特殊であるから、後述のシナリオ 11 で説明する)。

推奨

- (a) ST. 25 の配列表の<222>欄がある場合、ST. 26 のフォーマットに直接移行または別の方法で移行しても、主題事項が追加されるという問題は生じない。
- (b) ST. 25 に準拠する配列表に<222>欄がなく、出願明細書に位置情報が記載されている場合、ST.26 のフォーマットに直接移行または別の方法で移行しても、主題事項が追加されるという問題は生じない。
- (c) ST. 25 の配列表にも、出願明細書にも位置情報が記載されていない場合には、feature は配列全体に適用されると解釈する (location を配列全体より狭めてしまうと、出願明細書で裏付けなく示すことは、主題事項が追加または削除されてしまう可能性が高い。将来問題にならないように、元(ST. 25)の配列表および出願の明細書を作成する際は、可能な限り配列上の location を記載するよう注意すべきである)。

シナリオ 11

ST. 25 では、単一の連続したポリペプチドにコードされるが、イントロン等の1つまたは複数の数の非コード配列で区切られているコード配列を、以下に例示するように、複数の独立した feature「CDS」として示していた。

```
<220>
<221> CDS
<222> (1)..(571)
```

```
<220>
<221> CDS
<222> (639)..(859)
```

一方、ST. 26 では、指定された位置でコードされるポリペプチドが結合し、単一の連続したポリペプチドを形成することを示す location 演算子がある (NT. 25 および NT. 26 の両方とも、終止コドンが CDS の feature location に記載されていることが必要である)。

推奨

- (a) ST.25 の配列表または出願明細書において、複数の独立した feature「CDS」によってコードされたポリペプチド配列が単一の連続したポリペプチドを形成していることが明確に示されている場合、単一の feature「CDS」においてイントロンによって区切られたコード配列は、主題事項が追加されないように、以下に例示する存在位置演算子「join」を用いて示さなければならない。

```
<INSDFeature_key>CDS</INSDFeature_key>
<INSDFeature_location>join(1..571,639..859)</INSDFeature_location>
```

- (b) ST. 25 の配列表または出願明細書が、2つの別々の feature「CDS」によってコードされたポリペプチド配列が単一の連続したポリペプチドを形成していることを示していない場合、存在位置演算子「join」を用いると主題事項が追加される可能性があると思われる。

シナリオ 12

ST. 25 では、feature の名称は表 5 または表 6 に記載されている名称を用いなければならないと定めている。しかし米国の規定では、これらの名称は推奨されているが、必須ではないとしている。したがって ST. 25 の配列表(米国の規定に準拠)では、ST.26 に対応する feature key のない feature key「custom」の名称が設定される可能性がある。また、<221>欄に feature の名称が設定されていない場合や、<221>欄の記載が無い場合も考えられる。これらのシナリオは、同様の方法で処理することができる。

推奨

ST. 25 の feature key「custom」の名称は、ST.26 の配列表で、次のように主題事項を追加せずに表現することができる。

| 種類 | ST. 25 の feature key<221> | ST. 26 に準拠型の同等の feature key& qualifier | | |
|----|---------------------------|--|-----------|--|
| | | Feature key | qualifier | qualifier 値 |
| NA | feature key「custom」 | misc_feature | note | feature key「custom」の名称および、存在する場合は<223>の値 |
| AA | feature key「cutome」 | SITE または REGION | note | feature key「custom」の名称および、存在する場合は<223>の値 |

シナリオ 13

ST. 25 には feature key「VARSPPLIC」が記載されており、この feature key は「選択的スプライシングによって得られる配列の変異体についての記載」と定義されている。ST. 26 では、この feature key「VARSPPLIC」はより広義の feature key「VAR_SEQ」に置き換えられており、選択的スプライシング、選択的なプロモーターの利用、選択的な代替開始および、リボソームフレームシフトにより得られる配列の変異体の記載」と定義されている。したがって、ST. 26 の配列表では、feature key「VARSPPLIC」の代わりに feature key「VAR_SEQ」を、詳しい説明無しに使用すべきではない。

推奨

ST. 26 では、feature key「VAR_SEQ」を、qualifier「note」とともに使用すべきである。qualifier「note」の値には、ST. 25 の狭義の範囲の説明（例、「選択スプライシングにより得られる変異体」）が記載されるべきである。ST. 25 の <223>欄に記載されている他の情報も、qualifier「note」に記載すべきである。

シナリオ 14

配列が人工的に発生する場合は、ST. 25 の<213>「Organism」欄には「Artificial Sequence」（人工配列）との記載が必要である。ST. 26 では、feature key「source」は、qualifier「organism」を必要とし、qualifier 値には「Artificial Sequence」ではなく「synthetic construct」（合成構造）と示さなければならない。

推奨

ST. 26 の qualifier「organism」の値には、「synthetic construct」を用いなければならない。主題事項が削除されないように、ST. 25 の<223>欄に記載されている説明情報は、qualifier「note」（この場合の feature key は「source」）に記載されるべきである。

シナリオ 15

配列の元となる生物の学名が不明な場合、ST. 25 の<213> Organism 欄には「Unknown」と記載することが必要である。ST. 26 では、feature key「source」は、qualifier「organism」を必要とし、その qualifier の値は「Unknown」ではなく、「Unidentified」と表示されなければならない。

推奨

ST. 26 の qualifier「organism」の値は、「unidentified」でなければならない。主題事項が削除されないように、ST. 25 の<223>欄に記載されている説明情報はすべて、ST. 26 の qualifier「note」（この場合の feature key は「source」）の値として使用すべきである。

シナリオ 16

ST. 25 では、アミノ酸の数え上げにおいて、成熟蛋白質に先行するアミノ酸、例えばプレ配列、プロ配列、プレ-プロ配列およびシグナル配列などについて、1 番の次のアミノ酸から逆算して負の数を任意に記載することができる。ST.26 では、feature location に負の数を入れることはできない。

推奨

- (a) ST. 25 の配列表に、<221>とそれに付随する<222>欄で表現された feature が、例えば「PROPEP」および/または「CHAIN」のように、負の方向やおよび/あるいは正の方向の番号付けがある場合、ST. 26 の配列表では、例えば「PROPEP」および/または「CHAIN」のような適切な feature key を使用すべきである。また、<223>欄に情報が記載されている場合は、それとともに qualifier「note」を qualifier の値として使用しても良い。
- (b) ST.25 の配列表に、<221>とそれに付随する<222>欄で表現された feature はなかったが、出願明細書に負の方向および/または正の方向の番号付けがある場合、ST.26 の配列表では、例えば「PROPEP」および/または「CHAIN」のような適切な feature key を使用すべきである。それ以外の場合は、feature key「REGION」を使用しても良い。出願明細書に詳細な情報が記載されている場合は、それとともに qualifier「note」を qualifier の値として使用しても良い。
- (c) ST. 25 配列表および出願明細書のいずれにも、負の方向および/あるいは正の方向の番号付けを説明する情報がある場合、ST. 26 の配列表では、主題事項が削除されないように、ST. 25 の負の番号が付された領域にまたがる feature location を示す feature key「REGION」を使用すべきである。また、優先権主張出願の ST. 25 の配列表において、アミノ酸配列に負の番号が付されていることを示すために、qualifier「note」を使用すべきである。

シナリオ 17

ST.25 では、<300>から<313>欄への記載情報を規定している。ST.26 ではそのような情報を記載することを規定していない。

推奨

ST. 25 の<300>から<313>欄への記載情報が、対応する出願書類に記載されていない場合は、挿入すべきである。

シナリオ 18

ST. 25 では、標準的な遺伝コード表以外の遺伝コード表を用いて塩基配列のコード領域を翻訳する標準的な方法はなかったが、ST. 26 には、feature key「CDS」の qualifier「transl_table」があり、この qualifier を使用することで、その領域が標準的な遺伝コード表以外の遺伝コード表を用いて翻訳されていることを示すことができる。qualifier「transl_table」が使用されない場合、標準的な遺伝コード表の使用として扱われる。

推奨

- (a) ST. 25 の配列表または出願明細書に、標準的な遺伝コード表以外の遺伝コード表を使ってコード領域が翻訳されていることを明確に示している場合、qualifier「transl_table」(値は、遺伝コード表の適切な番号)を用いなければならない。この qualifier を用いない場合は、「標準的な遺伝コード表」の利用がデフォルトであると認識されているため、ST.26 の配列表に、ST. 25 の配列表または出願明細書に記載のある標準的な遺伝コード表以外の遺伝コード表が用いられているという情報を記載しなかった場合、主題事項が削除される可能性が高い。
- (b) ST. 25 の配列表または出願明細書に、標準的な遺伝コード表以外の遺伝コード表を用いてコード領域が翻訳されていることが記載されていなかった場合、qualifier「transl_table」を使用すべきではなく、qualifier の値「1」(すなわち標準的な遺伝コード表が使用されている)のみを使用すべきである。qualifier の値「1」以外の qualifier の値と共に qualifier「transl_table」を用いると、主題事項が追加または削除される可能性が高い。

シナリオ 19

ST. 25 では feature の location を示す標準的な方法を示していない。特に指定された残基または残基の範囲を超えて延びる部位または領域に含まれる feature (例. 開示された配列の片端または両端を超えて延びる塩基配列のコード領域が挙げられる)である。これに対して ST. 26 の feature 存在位置記述子は、記号「<」または「>」を使用し、それらの部位や領域の location を示す標準的な方法を示している。例えば、feature「CDS」の location は、開示された配列自体に終止コドンが含まれていなくても、その location を、例えば、「1...>321」のように記載することで、終止コドンを含まなければならない。

推奨

- (a) ST. 25 の配列表では、feature の location が配列表内で開示されている配列以上の長さに亘ることが明確に示されていないが、location が開示によって裏付けられているか、または配列自体から明らか(例、feature「CDS」の終止コドンが配列に含まれていない)である場合には、ST. 26 の配列表に記号「<」または「>」を使用しても良く、それが主題事項の追加となることは無い。
- (b) ST. 25 の配列表に、feature の location が配列を超えていることが明示されておらず、その location が開示によって裏付けられておらず、また配列自体からも明らかでない場合、主題事項を追加することなく ST. 26 に準拠することは不可能である。この場合、優先権出願および配列表自体が恐らく不完全であると言って間違いはないであろう。このような状況では、ST. 26 の配列表に記載されている feature の location は、先願に優先して適用されることはない。元(ST.25)の配列表および開示する出願書類について、feature の完全な情報を記載するよう注意すべきである。

シナリオ 20

ST.25 の附属書 I では、塩基配列に DNA セグメントと RNA セグメントの両方が含まれている場合、<212>の値は「DNA」とし、結合した DNA/RNA 分子を、<220>から<223>の feature section で詳細に記載するよう求めているが、記載を要する詳細情報の正確な性質は明確に規定されておらず、この規定への順守は徹底されていない。ST.26 の第 55 項では、結合した DNA/RNA 分子の各セグメント（ST.26 では内部的な整合性確保のため「セグメント」を「fragment」ではなく「segment」と記載）について、セグメントの location を含む feature key「misc_feature」と、セグメントが DNA か RNA かを示す qualifier「note」を使用して詳述することを求めている。

推奨

- (a) ST. 25の配列表が、<221>「misc_feature」、<222>で適切なlocation、およびセグメントがDNAかRNAであるかについて<223>で表示を用いて記載されている場合、DNAセグメントおよびRNAセグメントを1つまたは複数のfeatureとして記載している場合、その情報をST. 26に準拠型の配列表に移行する場合、DNAセグメントおよびRNAセグメントそれぞれに「misc_feature」を使用することで、主題事項の追加に関する問題は生じないはずである。
- (b) ST. 25 の配列表が、<221>「misc_feature」以外の feature key、<222>で適切な location、およびセグメントが DNA か RNA であるかについて<223>で表示を用いて記載されている場合、その情報を ST. 26 に準拠型の配列表に移行する場合、DNA セグメントおよび RNA セグメントそれぞれに「misc_feature」を使用し、元の<221>の feature key を値とする qualifier「note」を追加することで、主題事項の追加または削除に関する問題は生じないはずである。
- (c) ST. 25 の配列表が、<221>および<222>欄に関連しない<223>欄に各セグメントの同一性（DNA または RNA）および location を記載している場合（例：Artificial Sequence(人工配列)の説明）、DNA セグメントおよび RNA セグメントそれぞれに「misc_feature」を使用してその情報を ST.26 に準拠型の配列表に移行することは、主題事項の追加に関する問題は生じないはずである。
- (d) ST. 25 の配列表が、<221> misc_feature および DNA と RNA が結合した分子であることを示している<223>を使用して feature で分子を説明しているが、各セグメントの location の情報を記載していない場合であって、さらに、
- (i) 明細書に DNA セグメントおよび RNA セグメントそれぞれの location の情報が記載されている場合、その location を DNA セグメントおよび RNA セグメントそれぞれの「misc_feature」を使用して ST.26 に準拠型の配列表に移行することで、主題事項の追加に関する問題は生じないはずである。
 - (ii) 明細書に DNA セグメントおよび RNA セグメントそれぞれの location の情報が記載されていない場合は、主題事項を追加せずに ST. 26 を遵守することは不可能であろう。この場合、優先権出願および配列表は、それ自身、間違いなく不完全である。このような状況では、ST. 26 の配列表に feature の位置が記載されていても、先願に優先権は与えられない。元の(ST.25)の配列表および開示する出願書類には、完全な feature の情報を記載するよう注意すべきである。
- (e) ST. 25 の配列表が、misc_feature 以外の<221>の feature key および、その分子が結合した DNA/RNA 分子であることを示している<223>を使用した feature で分子を記載しているが、各セグメントの location の情報が記載されていない場合であって、さらに、
- (i) 明細書に DNA セグメントおよび RNA セグメントそれぞれの location の情報が記載されている場合、その location を DNA セグメントおよび RNA セグメントそれぞれの詳細情報を feature table および元の<221>の feature key を値とする qualifier「note」を追加して使用し、その情報を ST.26 に準拠型の配列表に移行することで、主題事項の追加や削除の問題は生じないはずである。
 - (ii) 明細書に DNA セグメントおよび RNA セグメントそれぞれの location の情報が記載されていない場合は、主題事項を追加せずに ST. 26 を遵守することは不可能であろう。この場合、優先権出願および配列表は、それ自身、間違いなく不完全である。このような状況では、ST. 26 の配列表に feature の位置が記載されていても、先願に優先権は与えられない。元の(ST.25)の配列表および開示する出願書類には、完全な feature の情報を記載するよう注意すべきである。

- (f) ST. 25 の配列表が、例えば「Artificial Sequence」（人工配列）の説明のように、<223>欄に結合された DNA/RNA 分子を記載していたが、各セグメントの feature key や location の情報が記載されていない場合であって、さらに、
- (i) 明細書に DNA セグメントおよび RNA セグメントそれぞれの location の情報が記載されている場合、DNA セグメントおよび RNA セグメントそれぞれの misc_feature を使用してその情報を ST.26 に準拠型の配列表に移行することで、主題事項の追加の問題は生じないはずである。
 - (ii) 明細書に DNA セグメントおよび RNA セグメントそれぞれの location の情報が記載されていない場合は、主題事項を追加せずに ST. 26 を遵守することは不可能であろう。この場合、優先権出願および配列表は、それ自身、間違いなく不完全である。このような状況では、ST. 26 に準拠する配列表に feature の location が記載されていても、先願に優先権は与えられない。元の(ST.25)の配列表および開示する出願書類には、完全な feature の情報を記載するよう注意すべきである。

[附属書 VII および本標準は以上]