



WIPO SEQUENCE 2.3.0 版

用户手册

本文件的目的是向用户提供说明，介绍如何使用 WIPO Sequence 桌面应用程序来执行基本操作。通常情况下，用户是希望提交含有序列表的专利申请的专利申请人或其代理人。

WIPO SEQUENCE 2.3.0 版

用户手册

目录

页次

1	简介	6
	概述.....	6
	系统要求.....	6
	安装	8
	Windows.....	8
	Linux.....	8
	OSX.....	8
	无提示安装	8
	卸载	9
	Windows.....	9
	Linux.....	9
	OSX.....	9
	自动更新过程	9
2	工具功能.....	11
3	工具概述.....	14
	工具主要内容	14
	视图	15

分节	15
叠加层	15
表格	16
面板视图	17
编辑面板	17
日期选择器	18
PDF 阅读器	19
键盘导航	19
项目主页视图	21
项目页面	22
项目详细信息	24
基本信息	24
帮助	25
首选项	1
4 工具功能	2
项目主页	2
创建项目	2
导入项目	3
导入序列列表	6
验证序列列表	10
删除项目	13
人员和组织	13
创建人员或组织	13
自定义生物体	14

创建自定义生物体.....	15
导出自定义生物体.....	15
导入自定义生物体.....	17
系统首选项.....	18
项目详细信息.....	21
打印项目.....	21
导出项目.....	22
从另一个项目导入信息.....	23
验证项目.....	25
生成序列表.....	28
常规信息.....	30
申请识别.....	30
优先权识别.....	31
申请人和发明人.....	32
发明名称.....	33
序列.....	34
创建序列.....	34
导入序列.....	37
插入序列.....	41
重排序列.....	43
批量编辑.....	44
编辑序列.....	49
特征.....	50
限定符.....	58

语种相关限定符.....	64
导入报告.....	69
导入报告表.....	70
变更数据表.....	71
显示序列表.....	71
5 文件格式.....	76
ST.25	76
mULTI-SEQUENCE (多序列)	76
Raw	77
FASTA	77
有一个序列的 FASTA 文件.....	77

1 简介

概述

WIPO Sequence 是一个桌面工具，使用户能够：

- (i) 创建/编辑符合产权组织标准 **ST.26** 的 **XML** 格式序列表
- (ii) 根据产权组织标准 **ST.26**，验证 **XML** 格式的序列表是否符合要求。

可以在以下位置找到产权组织标准 **ST.26**：

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

本文件介绍了作为申请人或申请人的代理人，如何使用 **WIPO Sequence**。该工具的功能列表见本文档第 2 节。

系统要求

WIPO Sequence 工具支持以下操作系统：

- Windows 10 版本 1803（32 位和 64 位版本）
- Linux: Ubuntu 版本 18.04 和 CentOS 7 版本 1804
- MacOS 版本 10.13（64 位版本）

除上述版本外，还支持以下操作系统：

- Windows 7 和更高版本（32 位和 64 位）
- Ubuntu 12.04 及更高版本
- MacOS 版本 10.9（64 位版本）

WIPO Sequence 工具要求的最低硬件规格如下：

- CPU: 1.6 GHz
- 内存: 4 Gb

- 可用硬盘：1 GB（存储序列信息可能需要额外的硬盘空间）
- 屏幕分辨率：1366x768

安装

Windows

WIPO SEQUENCE 为 32 位和 64 位版本的 Windows 提供单一安装文件。用户应遵循安装向导中显示的过程。

我们必须说明，如果使用 32 位版本，用户会遇到性能问题，或者工具无法正常工作。

用于存储项目信息和工具日志的数据库文件存储在以下位置：

```
C:\Users\\AppData\Roaming\ST26_authoring
```

更新或卸载应用程序时，不会删除这些文件。因此，如果重新安装应用程序，项目数据将保留。

Linux

WIPO Sequence 以“AppImage”文件（<https://appimage.org/>）提供，可在大多数 Linux 发行版（包括 CentOS 和 Ubuntu）上运行。要运行文件，用户可以双击文件或使用命令行来执行。

一开始，系统将提示用户是否选择在桌面上创建快捷方式。

OSX

WIPO Sequence 提供“dmg”文件，用于在 MacOS 64 位操作系统上安装。为了安装它，用户应双击该文件并按照向导进行操作。

用于存储项目信息和工具日志的数据库文件存储在以下位置：

```
/Users/<username>/Library/Application Support/ST26_authoring
```

更新或卸载应用程序时，不会删除这些文件。因此，如果重新安装应用程序，项目数据将保留。

无提示安装

WIPO Sequence 支持无提示安装，在安装时使用以下标记（.exe 安装程序文件）：

- **/S:** 开启无提示安装
- **/allusers:** 安装该工具，供台式机上的所有 Windows 用户使用（只有作为管理员用户登入时，才能开启安装）。

卸载

Windows

WIPO Sequence 提供卸载向导，可以在 Windows 中的“添加或删除程序”选项下启动。

要完全删除日志文件和用于存储项目信息的文件，必须删除以下文件夹：

```
C:\Users\\AppData\Roaming\ST26_authoring
```

Linux

删除 Linux “ApplImage” 文件，该应用程序即从计算机上卸载。此外，可以通过从以下位置删除桌面文件来删除菜单项：

```
$HOME/.local/share/applications/.
```

要完全删除日志文件和用于存储项目信息的文件，必须删除以下文件夹：

```
/Users/<username>/.config/ST26_authoring
```

OSX

该应用程序可以在 OSX 上的 Finder 中，从“应用程序”部分卸载。

要完全删除日志文件和用于存储项目信息的文件，必须删除以下文件夹：

```
/Users/<username>/Library/Application Support/ST26_authoring
```

自动更新过程

启动 WIPO Sequence 后，该工具将检查是否有可用的更新，如果找到更新，它将要求用户升级至较新的稳定版本。

建议不要跳过更新，以确保使用的版本符合 ST.26 的最新版本。

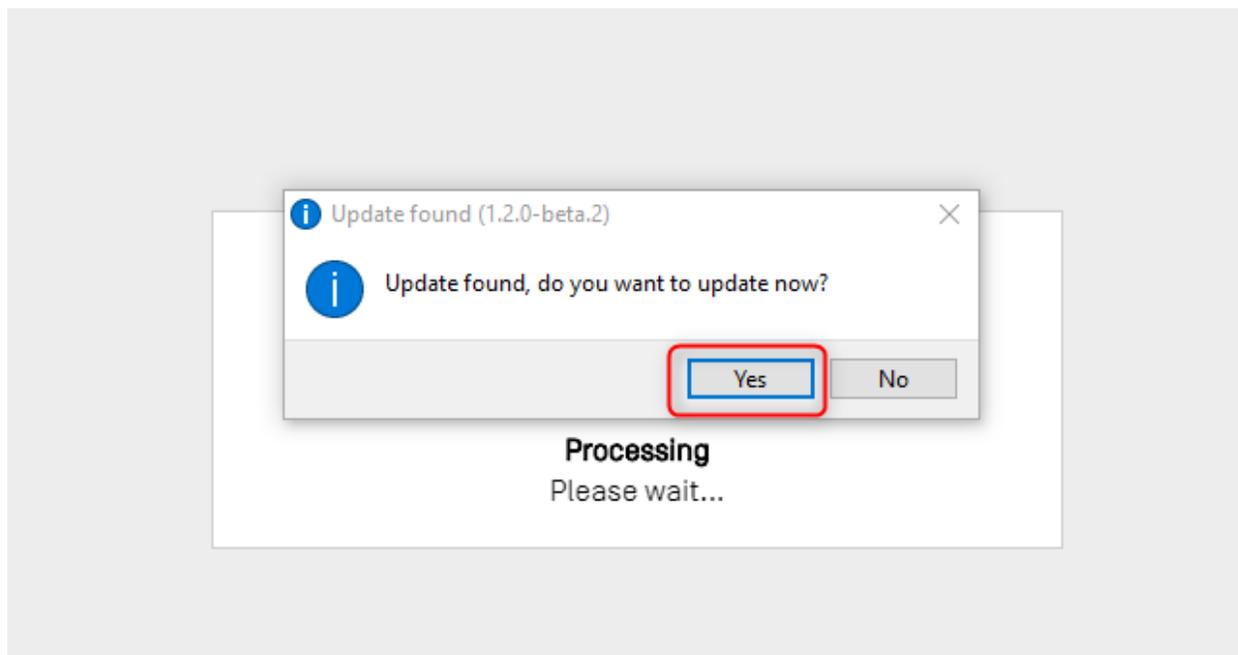


图 1：自动更新弹窗

注：

为启用自动更新，计算机必须连接到互联网，并且该工具必须有通过 HTTP 访问产权组织网站的权限。同时请注意，不会将用户的任何信息传送到产权组织更新服务器。

用户应耐心等待新版本的下载和安装，而不转移对应用的注意力。

2 工具功能

本节概述该工具当前版本实现的所有功能。

类别	功能
项目	创建项目，在其中存储与一个序列列表相关的数据
项目	生成序列列表
项目	编辑项目的属性
项目	导出自由文本限定符，用于以 XLIFF 格式进行翻译
项目	导出存储在项目中的所有数据，以便以后可以将其导入到系统的相同或不同实例中（项目元数据除外）
项目	以人类可读格式（.html 和.txt）显示/导出生成的序列列表
项目	将存储在项目文件（.zip）中的所有数据导入到新创建的项目中
项目	将数据从 ST.25 序列列表文件导入到新创建的项目中
项目	将数据从 ST.26 序列列表文件导入到新创建的项目中
项目	从 FASTA 序列中将数据导入已有项目
项目	从 ‘multi-sequence’ 格式文件导入序列数据
项目	将另一个项目（原始项目）的数据导入当前项目（目标项目）
项目	从项目打印数据
项目	从生成的 ST.26 序列列表打印数据
项目	在报告中记录在导入时已更改的数据，注明原始数据和更改的数据
项目	验证 ST.26 序列列表文件，并将问题作为包含警告和错误消息的验证报告列出

项目	验证存储在项目中的数据，并将问题作为包含警告和错误消息的验证报告列出
项目	删除项目
常规信息	将发明名称及其相应的语言代码添加到项目中
常规信息	将申请信息（当前或先前的申请）添加到项目中
常规信息	将申请人或发明人信息添加到项目中
序列	将源特征及其强制限定符添加到序列中
序列	将特征信息添加到序列中
序列	将限定符信息添加到特征
序列	在项目中创建序列
序列	批量编辑或添加特征到一定范围的序列表
序列	批量跳过一定范围的序列表
序列	批量删除一定范围的序列表
序列	在表的另一个位置创建和/或插入一个序列
序列	将序列导入到已有项目中
序列	删除序列
序列	为选定的 CDS 特征自动创建翻译限定符，并选择性地创建由 protein_id 识别的相关氨基酸序列
序列	编辑限定符数据
序列	编辑序列数据
序列	编辑特征键数据
序列	指明选定特征的位置信息

序列	检验残基不包含无效符号
序列	将氨基酸序列的分子属性设置为预定义值之一
序列	根据指定的遗传密码表号码（默认为遗传密码 1）对核酸序列进行翻译
序列	以 XLIFF 格式导入自由文本限定符源语言-目标语言对，以便在项目中提供翻译
序列	批量编辑序列注释，包括限定符 mol_type
序列	使用批量编辑批量删除一定范围的序列
序列	对序列表中的序列重新排序
自定义生物体	将新的生物体名称添加到此系统中存储的生物体名称列表中
自定义生物体	将自定义生物体名称列表导出到一个文本文件，该文本文件以后可以导入到 WIPO Sequence 的不同实例中
自定义生物体	从 XLIFF 文件将自定义生物体名称列表导入到已有项目中
人员/组织	在系统中存储有关申请人或发明人的信息（例如：姓名，对应的语言代码及其译文或拉丁字符音译（如适用），地址等），以便以后可在各种项目中使用
人员/组织	创建新的人名或组织名称
系统首选项	调整工具的某些首选项，影响该实例生成的所有项目

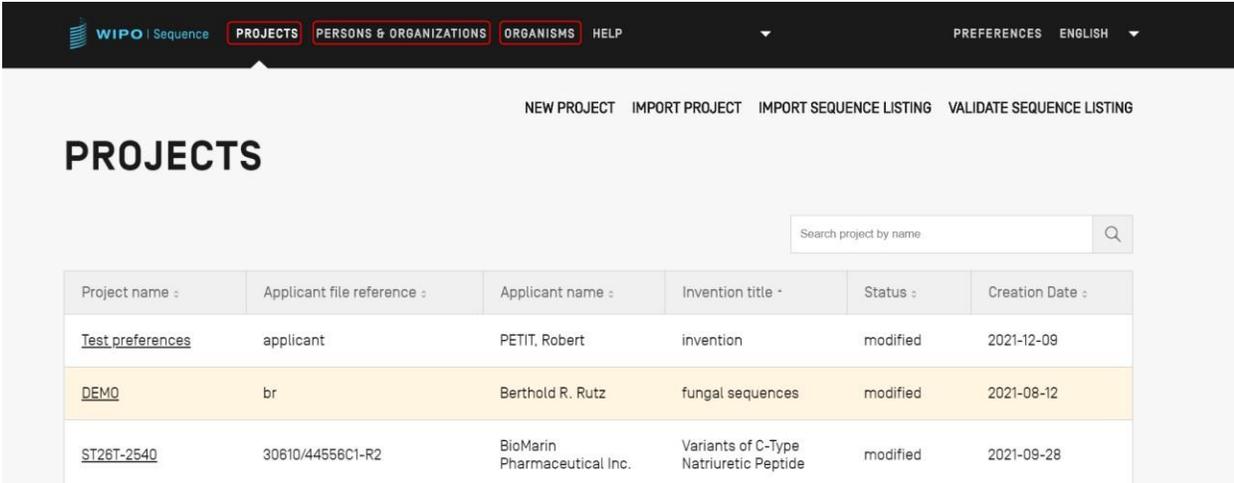
3 工具概述

工具主要内容

本节详细介绍了该工具的主要元素。提供此详细信息的目的是使用户熟悉桌面工具中提供的常见组件。

页面是该工具内视图的主要容器。有两个主要的页面，其中包含视图：

- 主页面
 - 项目主页视图
 - 人员和组织视图
 - 自定义生物体视图
 - 首选项视图
- 项目页面（可从项目视图访问）
 - 项目详细信息视图
 - 验证报告视图
 - 语种相关限定符视图
 - 导入报告视图
 - 显示序列列表视图



Project name :	Applicant file reference :	Applicant name :	Invention title :	Status :	Creation Date :
Test_preferences	applicant	PETIT, Robert	invention	modified	2021-12-09
DEMO	br	Berthold R. Rutz	fungal sequences	modified	2021-08-12
ST26T-2540	30610/44556C1-R2	BioMarin Pharmaceutical Inc.	Variants of C-Type Natriuretic Peptide	modified	2021-09-28

图 1：项目主页视图

视图

可以在同一页面中看到的的信息的不同显示称为“视图”。

分节

有些视图可以有分节。分节提供了一种方便的方式来分隔大视图的不同部分。

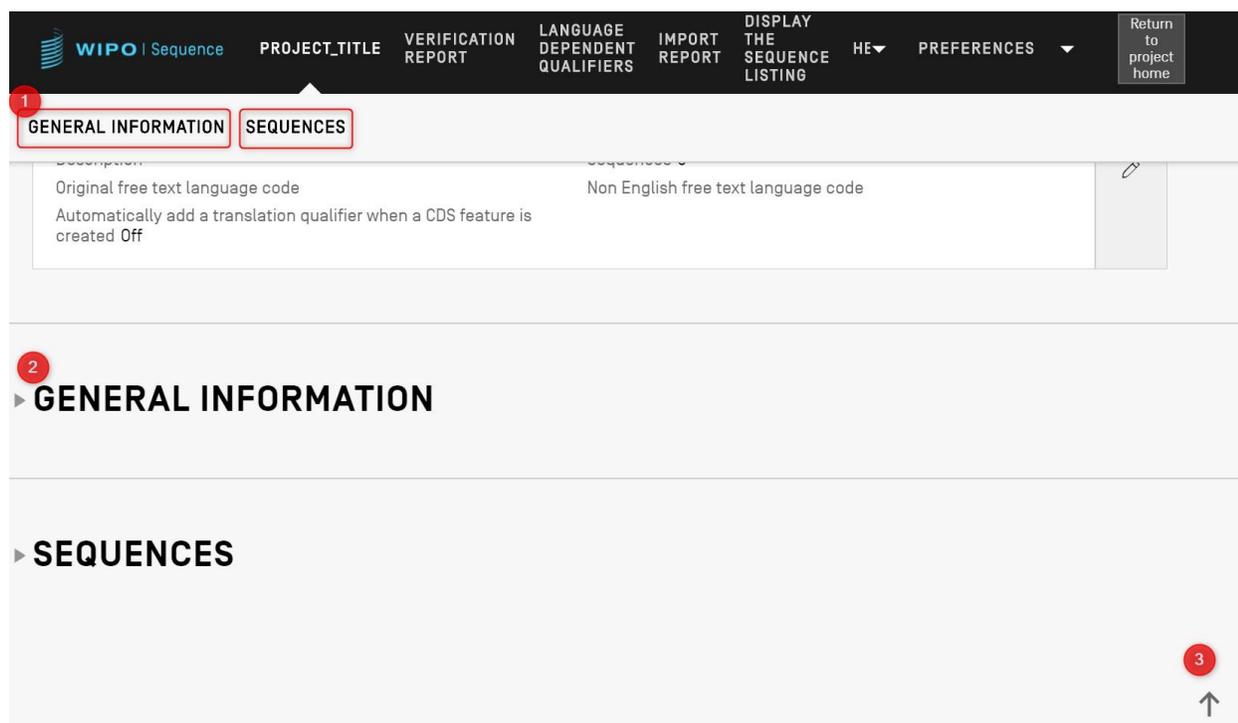


图 2：项目详细信息视图

在项目详细信息视图中，如图 1 所示，有两个可折叠节：常规信息和序列。在项目详细信息视图的顶部，有两个链接可将用户导航到相应的节（1）；为便于导航，每个节都可折叠（由左侧的小三角表示）（2）；用户可以单击箭头图标滚动到项目详细信息视图的顶部（3）。

叠加层

当需要填写或修改面板时，有时会在当前视图上方显示一个叠加层，使背景变灰，如错误!未找到引用源。中捕获的例子所示。

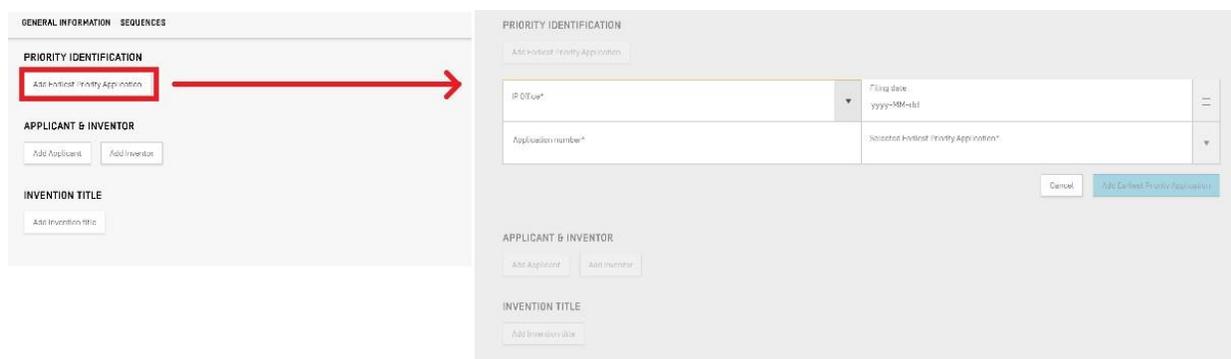


图 4: 叠加层

当叠加层可见时，叠加层后面的所有元素均不起作用，只有焦点所在的项目可被修改。

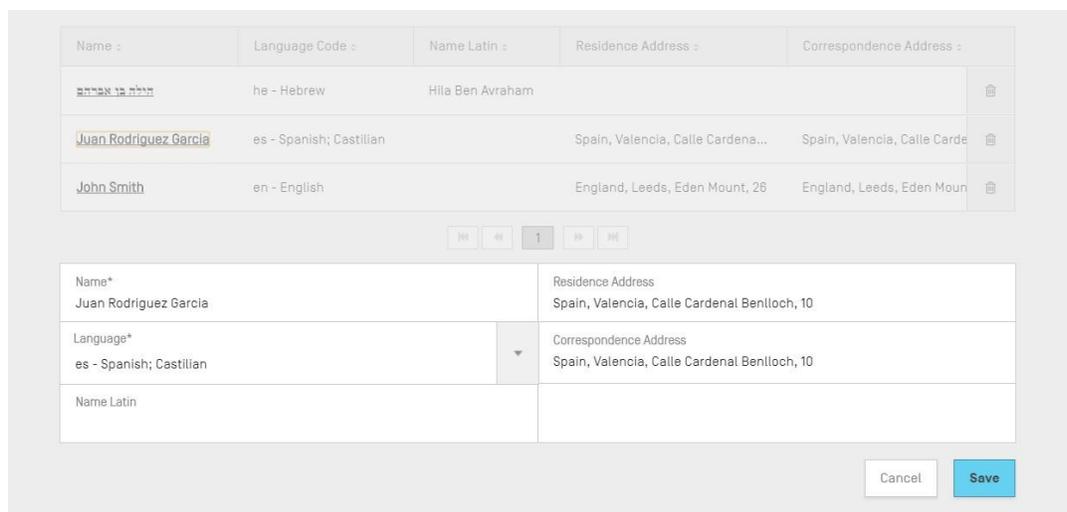
表格

Name	Language Code	Name Latin	Residence Address	Correspondence Address
<u>הילה בן אברהם</u>	he - Hebrew	Hila Ben Avraham		
Juan Rodriguez Garcia	es - Spanish; Castilian		Spain, Valencia, Calle Cardenal ...	Spain, Valencia, Calle Cardenal ...
<u>John Smith</u>	en - English		England, Leeds, Eden Mount, 26	England, Leeds, Eden Mount

图 5: 表格样例

表格样例显示在图 5。为表格启用搜索后，用户可以输入一些搜索数据，然后单击搜索栏 (1) 右侧的搜索图标。要返回表内元素的完整列表，请清除搜索输入框，单击搜索图标 (1)。

如果所有的条目不能放在一页中，用户可以通过单击表格下方的按钮在表格的行页面中导航 (2)。要删除表格中的条目，用户可以单击相应行右侧的回收站图标 (3)。要打开表格内条目的编辑视图，用户必须单击行内带下划线的值（下划线表示可以点击） (4)。



Name	Language Code	Name Latin	Residence Address	Correspondence Address
הילה בן אברהם	he - Hebrew	Hila Ben Avraham		
Juan Rodriguez Garcia	es - Spanish; Castilian		Spain, Valencia, Calle Cardena...	Spain, Valencia, Calle Carde
John Smith	en - English		England, Leeds, Eden Mount, 26	England, Leeds, Eden Moun

Name*	Residence Address
Juan Rodriguez Garcia	Spain, Valencia, Calle Cardenal Benlloch, 10
Language*	Correspondence Address
es - Spanish; Castilian	Spain, Valencia, Calle Cardenal Benlloch, 10
Name Latin	

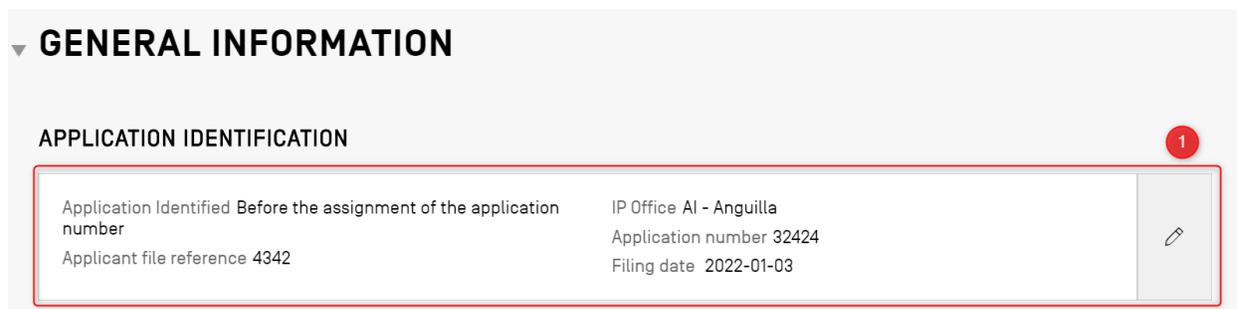
Cancel Save

图 6：编辑表格

选择此条目后，屏幕上会出现一个叠加层，表格下方会打开一个可编辑的面板。

面板视图

面板视图显示一组数据。面板视图中的项目可能分布在几列中。每个项目都有一个标签和一个（可选）值。



▼ **GENERAL INFORMATION**

APPLICATION IDENTIFICATION

Application Identified Before the assignment of the application number	IP Office AI - Anguilla
Applicant file reference 4342	Application number 32424
	Filing date 2022-01-03

1

图 7：面板视图样例

如图所示，在面板视图上单击带有铅笔图标（1）的按钮时，面板视图会被编辑面板所取代。

编辑面板

打开“编辑面板”后，将向用户显示可以修改的字段。用户完成对值的编辑后，将能够保存所做的更改，或者通过分别单击“保存”（1）或“取消”（2）按钮来放弃更改（如图 3：所示）。

注：

仅用于显示且不可修改的字段将显示为灰色。

▼ GENERAL INFORMATION

APPLICATION IDENTIFICATION

IP Office GB	
Application Identification* <input checked="" type="radio"/> Before the assignment of the application number <input type="radio"/> After the assignment of the application number	Applicant file reference* ABC1234 <small>Application File Reference is a mandatory field when 'Before the assignment of the application number' is provided.</small>
Application number 98968268463829	Filing date 2019-02-21

Cancel Save

图 3：编辑面板

日期选择器

每当需要用户输入日期时，就会出现日期选择器界面。通过单击日历图标（1），日期选择器组件将打开（2），如图 4 图 4 所示。

Year: 2019, Month: Feb

Calendar grid showing dates from 27 to 02 for the next month.

Filing date: 2019-02-21

图 4：日期选择器

当前选择的日期用灰色背景标记（22 日），使用日期用黑色下划线表示（26 日）。星期日视为每周的第一天。

反之，用户也可以按照适当的格式（“YYYY-MM-DD”）直接输入所需的日期（3）。

PDF 阅读器

当用户为特定序列列表、导入报告或项目信息打印验证报告时，将生成 PDF 文件并在 PDF 阅读器中打开。要下载并保存该文件，用户必须单击查看器（1）右上角的下载图标，如图 510 所示。



图 5: PDF 查看器

键盘导航

WIPO Sequence 支持基本的键盘导航。“TAB”键用于在不同项目之间跳转，“SPACE”键用于选择复选框和单选按钮。

图 11 中导航过程中的焦点是可见的：

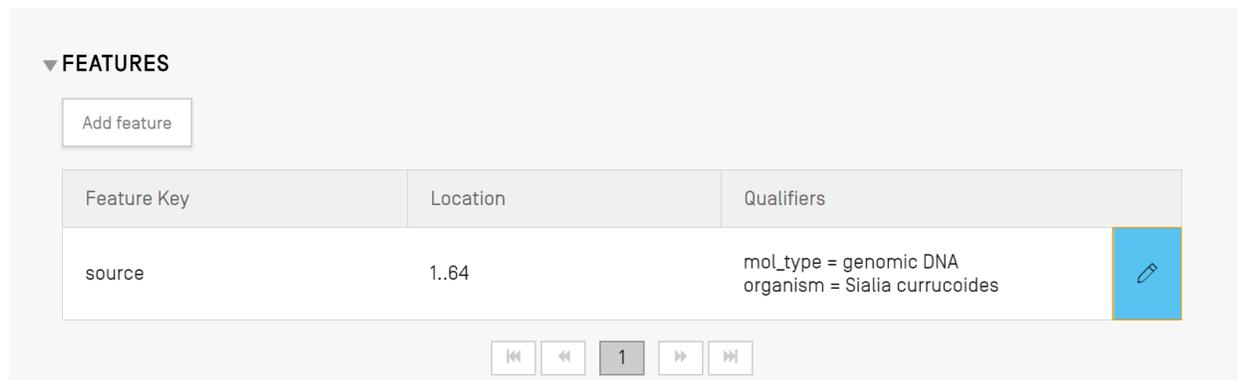


图 6: 编辑功能按钮的焦点

此外，将光标和焦点设置到左上角的输入字段。

例如，创建新项目时只有一个必填字段：名称。光标和焦点已经设置到项目名称字段，如下方图 7 所示：

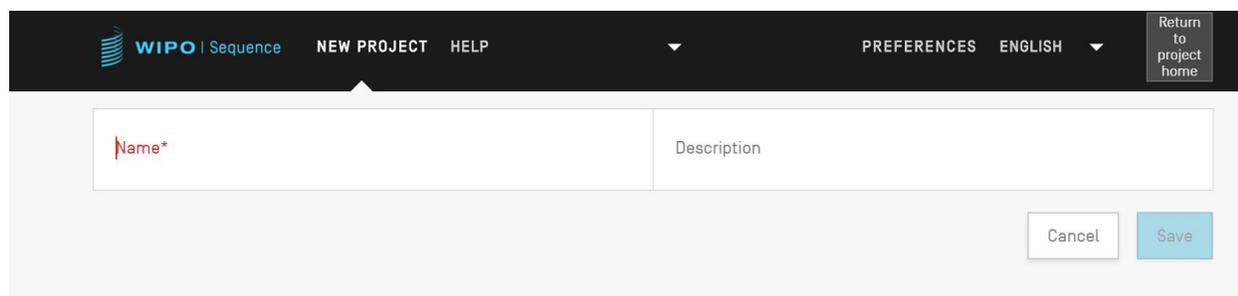


图 7: 光标聚焦在必填字段

项目主页视图

该工具的项目主页视图包含三个主要视图：

项目视图（见图 8）、**人员和组织视图**（见图 94）以及**生物体视图**（见图 105）。另外还有首选项（见图 19，位于右上角，适用于所有项目）。

Project name :	Applicant file reference :	Applicant name :	Invention title :	Status :	Creation Date :
cds_feature	15123-WO-PCT[2]	Novozymes A/S	Polypeptides Having Beta-Glucanase Activity and Polynucleotides Encoding Same	modified	2021-09-07
160K.500.SEQ	A400:66076	University of Tokyo	COPOLYMER INCLUDING UNCHARGED HYDROPHILIC BLOCK AND CATIONIC POLYAMINO ACID BLOCK HAVING HYDROPHOBIC GROUP IN PART OF SIDE CHAINS, AND USE THEREOF	modified	2021-09-07
160K.100.SEQS	A400:66076	University of Tokyo	aaaaaaaa	modified	2021-09-06
cds_feature	15123-WO-PCT[2]	Novozymes A/S	Polypeptides Having Beta-Glucanase Activity and Polynucleotides Encoding Same	invalid	2021-09-02
cdsFeatures	15123-WO-PCT[2]	Novozymes A/S	Polypeptides Having Beta-Glucanase Activity and Polynucleotides Encoding Same	modified	2021-09-01

图 8：项目主页视图

Name :	Language Code :	Name Latin :	Residence Address :	Correspondence Address :
Джо Смит				
Джейн Эйр				
Влад	ru - Russian	test name		

图 9：人员和组织视图

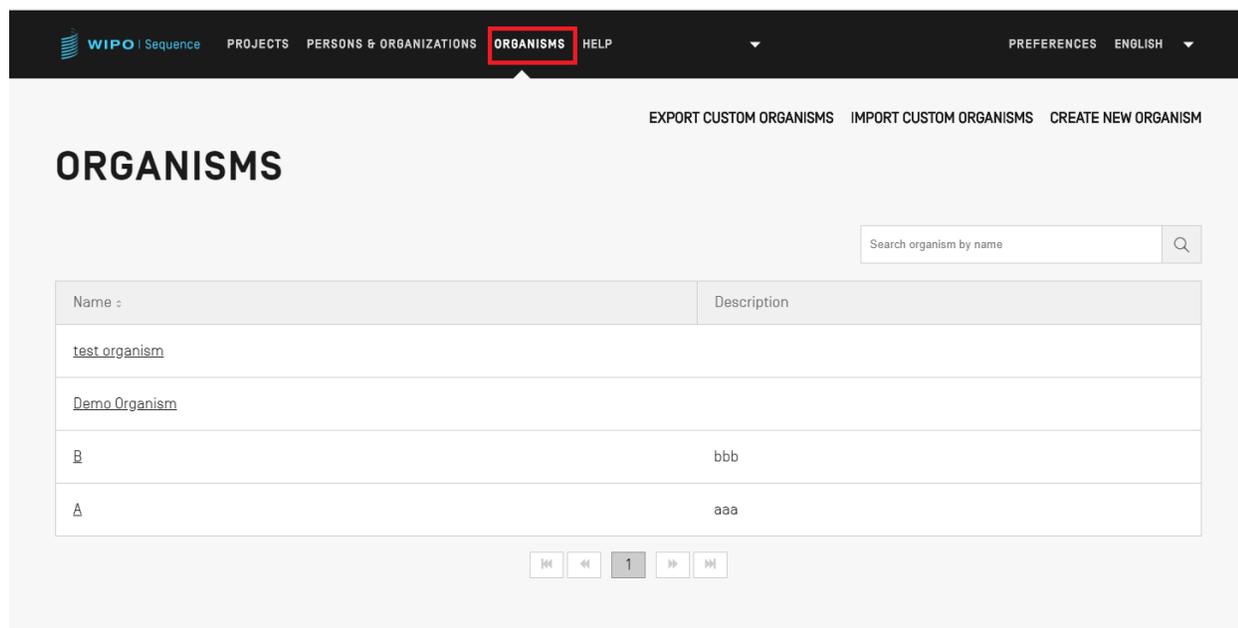


图 10 生物体视图

项目页面

项目页面由六个视图组成，这六个视图可用于在工作流的不同部分之间导航，如图 43 所示：

1. 项目详细信息视图 (1) (用项目名称表示，显示为“项目”)：包含所有项目数据的主视图，
2. 验证报告视图 (2)：可获取验证报告，
3. 语种相关限定符视图 (3)：可获取和导出/导入语种相关自由文本限定符，
4. 导入报告视图 (4)：可获取导入报告，
5. 显示序列列表视图 (5)：可获取生成的 ST.26 序列列表的人类可读格式，
6. 帮助菜单：包括对用户手册、产权组织标准 ST.26、WIPO Sequence 和 ST.26 知识库的引用 (6)，以及
7. 首选项视图 (7)。

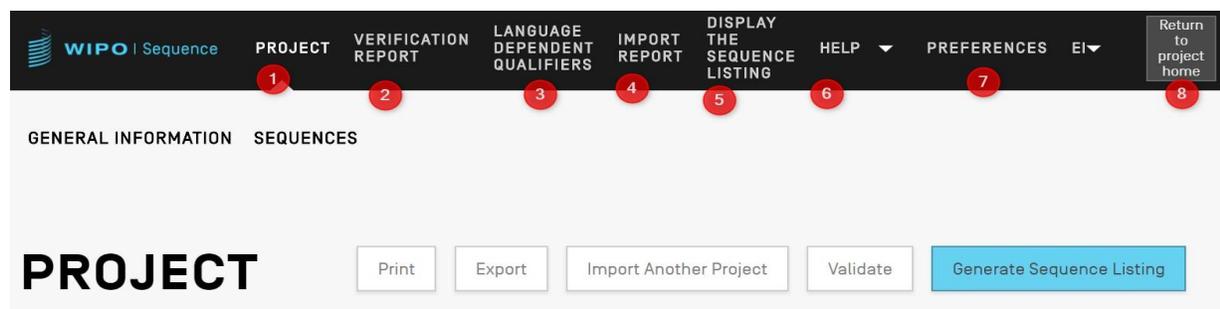


图 16: 11 项目详细信息工具栏

要返回项目主页，用户可以点击页眉/工具条右端的“返回项目主页”按钮（8）。

项目详细信息

基本信息

在项目详细信息视图的顶部可以找到一个包含项目基本信息的表格，如图 12 所示。

本节包括：

- 项目名称
- 项目创建日期和时间
- 对项目进行上一次更新的日期和时间。
- 项目状态（可能值：‘new’/‘modified’/‘generated’/‘invalid’/‘valid’/‘warnings’）-注意这不是可编辑的字段
- 项目说明-可选
- 所导入文件的名称（在导入项目的情况下）
- 自由文本限定符原始自由文本语言代码
- 序列数目（标签为：‘Sequences’）
- 创建 CDS 特征时，可调用自动添加翻译限定符的复选框（项目级功能）
- 非英语自由文本语言代码

Project Name	ST26T-2136	Creation date	2023-04-06 10:45
Last modified	2023-04-06 16:39	Status	generated
Description		File Name	ST26T-2136_3
Original free text language code		Sequences	11
Automatically add a translation qualifier when a CDS feature is created	Off	Non English free text language code	

图 12: 基本信息节

帮助

该工具将使用户能够查看可以使用的帮助选项，这有助于更好地使用这一工具（见图 13）。

帮助选项定向至以下信息：

- 提供用户手册的链接¹

¹出版时指向用户手册英文版，离线提供或在线提供其余 9 种 PCT 语言

- 提供 ST.26 知识库的链接²
- 提供联系 WIPO Sequence 团队联系表的链接
- 提供产权组织 ST.26 的链接³
- 提供 WIPO Sequence 桌面工具的基本信息

²出版时指向 WIPO Sequence ST.26 知识库

³出版时指向英文、法文或西班牙文版在线最新版本的标准

注：

与 WIPO Sequence 最新版配套的是用户手册和 WIPO ST.26，只有英文版。当用户为图形用户界面选择不同语言时，需要连接至互联网，才能获取其他语言版本的标准和用户手册。但是，提供的链接将指向与其选择相匹配的语言版本。

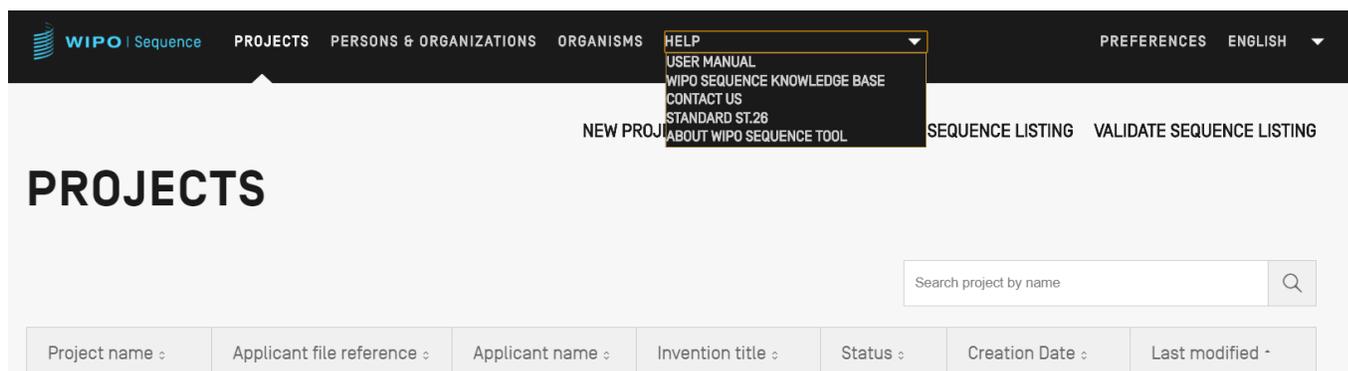


图 13: 帮助下拉菜单

首选项

在最后一个视图中，用户可以设置适用于所有项目的具体属性（见图 14）。关于这些参数中每个参数的具体详情可在下文的相关“工具功能”节找到。

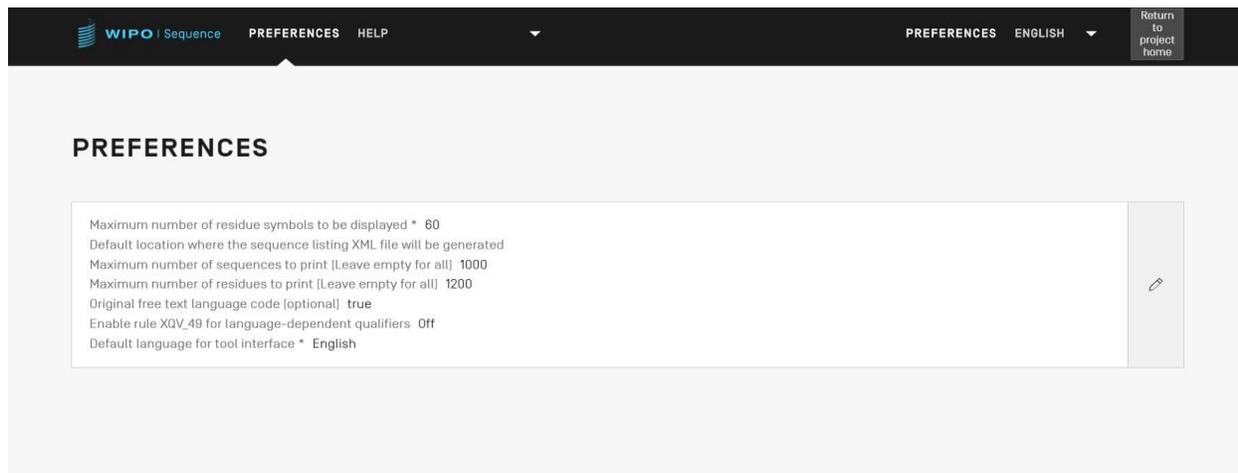


图 14: 首选项视图

4 工具功能

项目主页

本节详细介绍了在“项目主页视图”中可访问的不同功能。

项目是该工具用来存储生成序列列表所需数据的对象结构。一旦确认项目中存储的数据符合产权组织标准 **ST.26**，该工具就会使用这些数据作为生成的序列列表中的值。

在该视图中，显示已创建的项目列表，用户可以选择排序或使用检索功能，按项目名称、申请人文件号、申请人姓名、发明名称、状态或创建日期进行筛选。

注：

该工具最多可显示 1000 个项目。如果某个项目未在项目主页视图显示，则用户应使用检索功能，通过其名称来找到该项目，因为项目仍存储在本地，只是在该视图中不可见。

创建项目

要创建一个新项目，用户必须从项目主页视图开始（参见图 15）。

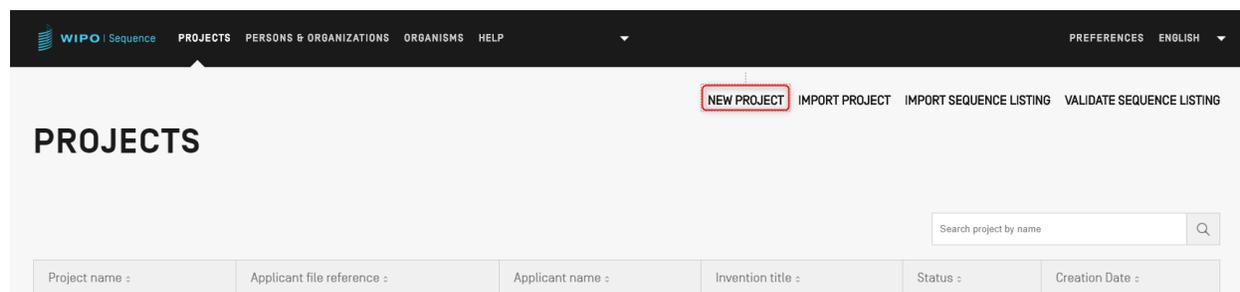
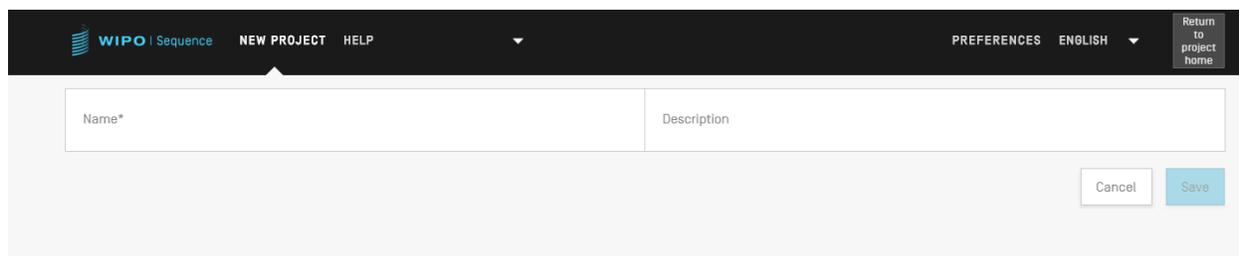


图 15: 创建新项目第 1 步

- 1) 单击图 15 所示视图顶部的“NEW PROJECT”链接。在接下来的视图中，该工具将要求提供‘Name’（必填）和‘Description’（选填）。



WIPO Sequence NEW PROJECT HELP

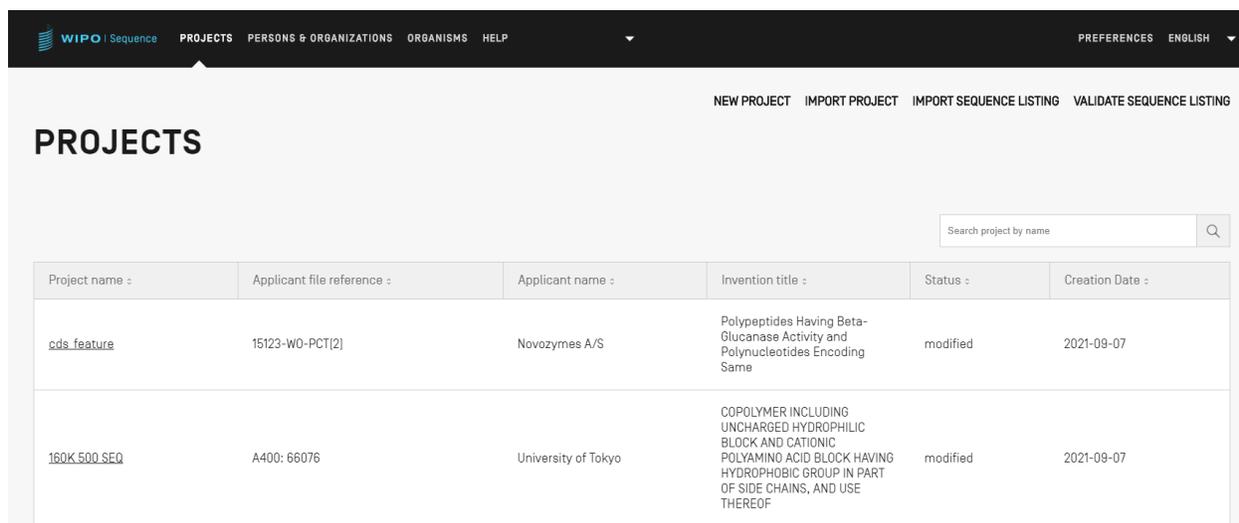
PREFERENCES ENGLISH Return to project home

Name* Description

Cancel Save

图 16: 创建新项目第 2 步

- 在名称字段中输入值时，将启用“Save”按钮，以使用户保存新项目。项目主页视图中包含这个新项目的项目列表如 17 所示。



WIPO Sequence PROJECTS PERSONS & ORGANIZATIONS ORGANISMS HELP

PREFERENCES ENGLISH

NEW PROJECT IMPORT PROJECT IMPORT SEQUENCE LISTING VALIDATE SEQUENCE LISTING

PROJECTS

Search project by name

Project name :	Applicant file reference :	Applicant name :	Invention title :	Status :	Creation Date :
cds_feature	15123-WO-PCT[2]	Novozymes A/S	Polypeptides Having Beta-Glucanase Activity and Polynucleotides Encoding Same	modified	2021-09-07
160K_500_SEQ	A400: 66076	University of Tokyo	COPOLYMER INCLUDING UNCHARGED HYDROPHILIC BLOCK AND CATIONIC POLYAMINO ACID BLOCK HAVING HYDROPHOBIC GROUP IN PART OF SIDE CHAINS, AND USE THEREOF	modified	2021-09-07

图 17: 以摘要形式显示的新项目

导入项目

此功能允许将以前导出的项目导入到该工具中，如**错误!未找到引用源。**所示。要导入项目文件，用户必须从项目主页视图开始，如下**错误!未找到引用源。**所示。

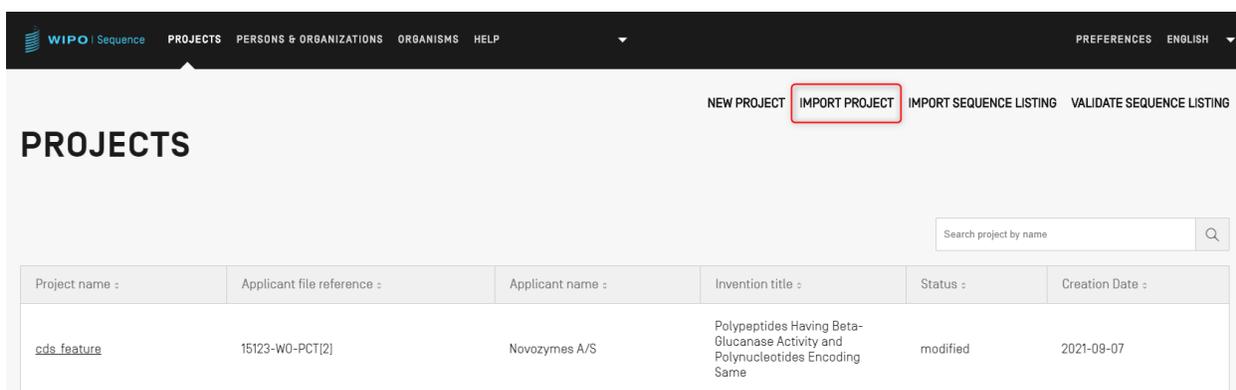


图 18: 导入项目第 1 步

1) 单击视图顶部的“IMPORT PROJECT”链接，如图 19: 所示。

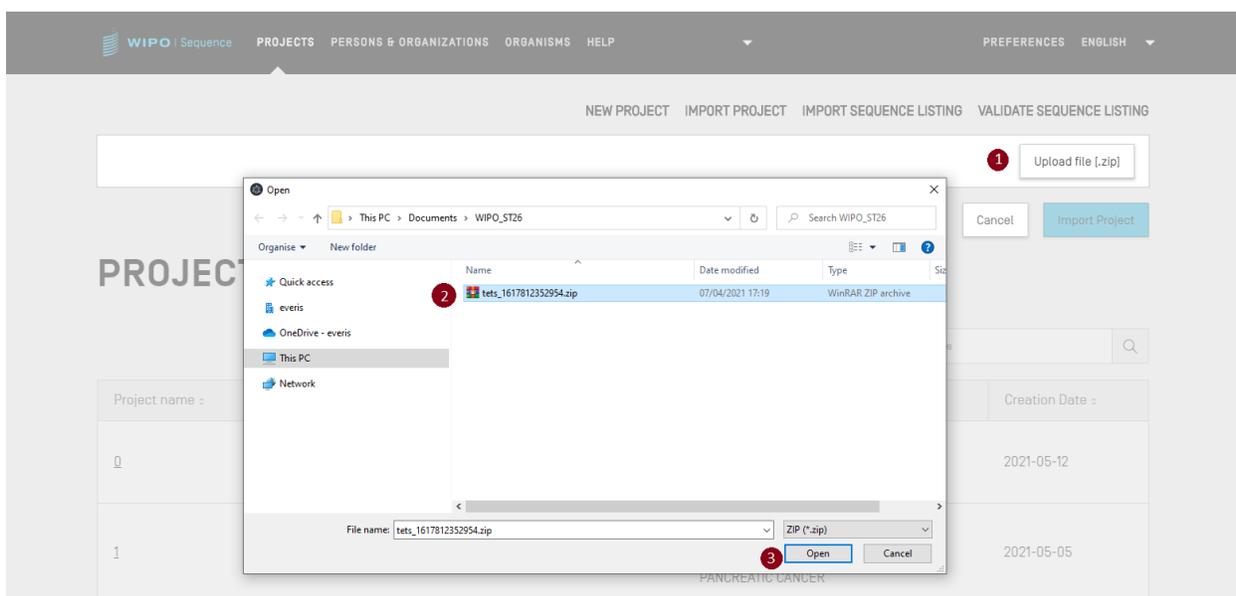


图 19: 导入项目对话框

- 2) 在错误!未找到引用源。所示的叠加层屏幕上，单击“Upload file [.zip]”按钮 (1)。
- 3) 在错误!未找到引用源。所示的可打开的对话框中，选择要导入的项目文件 (2 和 3)。

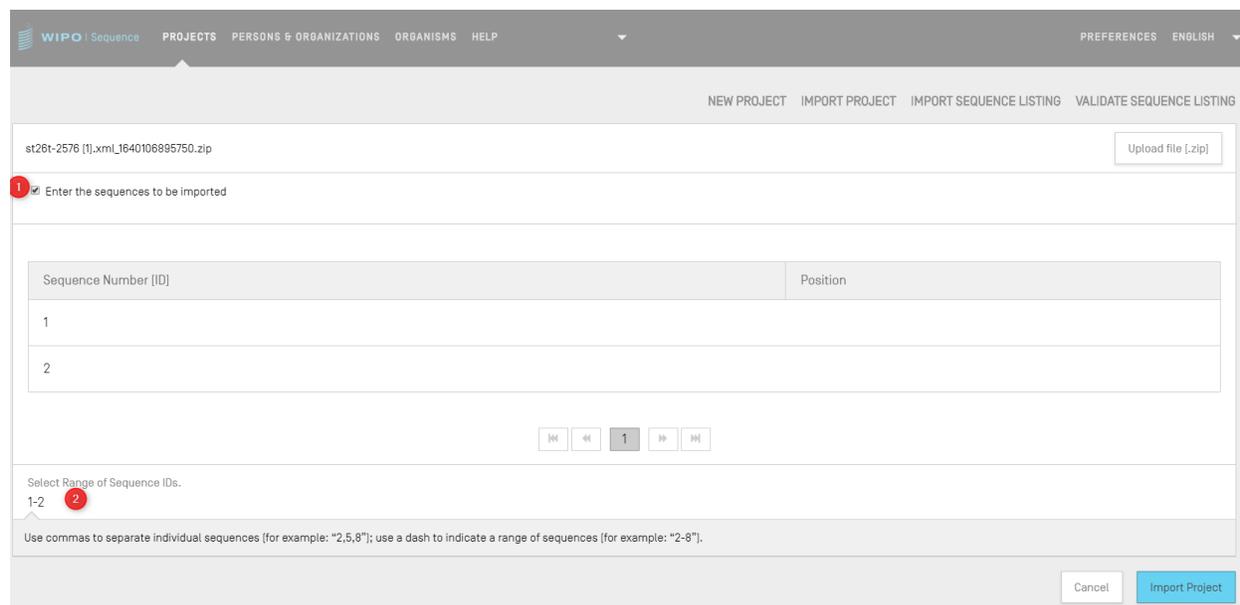


图 20: 导入项目，选择序列范围

- 4) 如果“**Select Range Sequences**”仍未选中，则将导入所有序列。如果用户希望对导入项目中的序列有所选择，则勾选“**Select Range Sequences**”复选框（1），并在适当字段（2）中输入所需序列的 ID 号，如图 21 所示。可以输入单个序列，也可以输入用逗号分隔的序列列表或 x-y 形式的序列范围。

示例：“1, 3, 7, 13-20, 30-50”

默认情况下，导入项目的总序列数将显示为一个范围，即：1-总序列数

- 5) 最后一步是单击蓝色的“导入项目”按钮（3），如图 21 所示。

如果项目成功导入，在视图的顶部会出现以下蓝色横幅和消息，如**错误!未找到引用源。**所示。

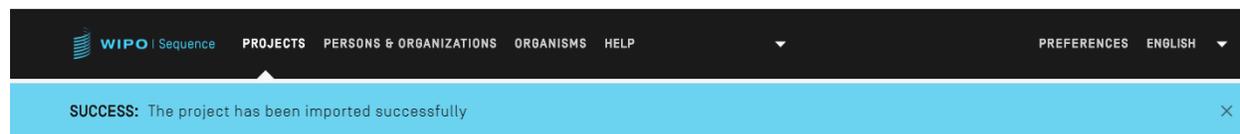


图 21: 项目成功导入，蓝色横幅

已知问题：WIPO Sequence 在导入项目时会产生很多意外错误：不清楚导致这种错误的原因。这大多数时候是由于挂起的工具而产生，如遇到问题，请尝试再次进行导入过程。

导入序列表

从项目主页视图中，用户可以专门从符合 ST.26 或 ST.25 的序列表中导入序列信息。ST.26 格式的文件格式为*.xml，ST.25 格式的文件格式为*.txt。

- 1) 首先，单击图 22：所示视图顶部的“IMPORT SEQUENCE LISTING”按钮。



图 22：导入序列表第 1 步

- 2) 在错误!未找到引用源。所示的打开的叠加层中，单击“Upload file ST.25 [.txt] or ST.26 [.xml]”按钮 (1)，在打开的对话框中选择所需的序列表文件 (2)。然后在“Project Name”字段中给出的名称下输入要创建的新项目的名称 (3)。

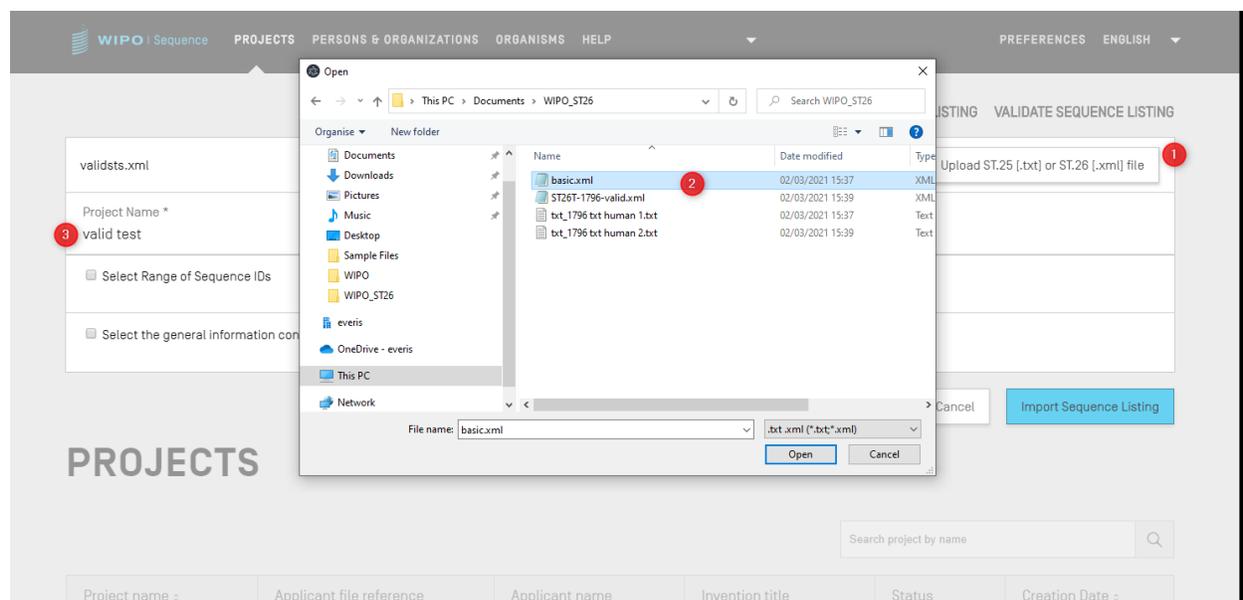


图 23：导入序列表，对话框

图 24 所示的两个主要复选框——“Select Range Sequences”和“Select the general information contents to be imported”，将允许用户指明将哪些部分导入到新项目中。

st26t-2624.xml Upload ST.25 [.txt] or ST.26 [.xml] file

Project Name *

Enter the sequences to be imported

Sequence Number [ID]	Position
[1]	1

1

Select Range of Sequence IDs.
1

Use commas to separate individual sequences [for example: "2,5,8"]; use a dash to indicate a range of sequences [for example: "2-8"].

Select the general information contents to be imported

Cancel Import Sequence Listing

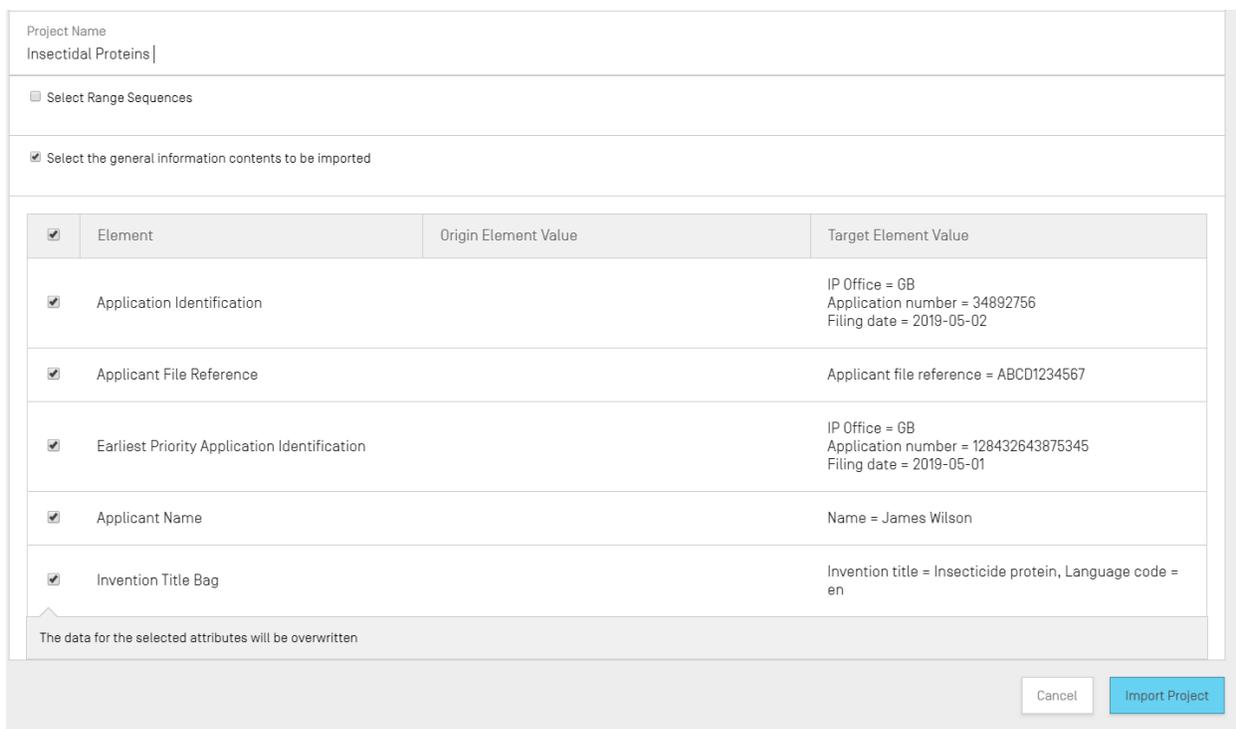
图 249：导入序列表，选择范围

第一个复选框允许用户输入希望从序列表中要导入的特定序列。可以输入单个序列，也可以输入用逗号分隔的序列表，或者输入 **x-y** 形式的序列范围。

示例：“1, 3, 7, 13-20, 30-50”

默认情况下，导入序列表的序列总数将显示为一个范围。

图 24 所示的表格有两列，一列给出了序列 ID 号，用于识别相应的序列，另一列给出它将在序列表中出现的“Position”。



The screenshot shows a web interface for importing sequence data. At the top, the 'Project Name' is 'Insectidal Proteins'. Below this, there are two checkboxes: 'Select Range Sequences' (unchecked) and 'Select the general information contents to be imported' (checked). A table lists various attributes with their original and target values. At the bottom right, there are 'Cancel' and 'Import Project' buttons.

<input checked="" type="checkbox"/>	Element	Origin Element Value	Target Element Value
<input checked="" type="checkbox"/>	Application Identification		IP Office = GB Application number = 34892756 Filing date = 2019-05-02
<input checked="" type="checkbox"/>	Applicant File Reference		Applicant file reference = ABCD1234567
<input checked="" type="checkbox"/>	Earliest Priority Application Identification		IP Office = GB Application number = 128432643875345 Filing date = 2019-05-01
<input checked="" type="checkbox"/>	Applicant Name		Name = James Wilson
<input checked="" type="checkbox"/>	Invention Title Bag		Invention title = Insecticide protein, Language code = en

The data for the selected attributes will be overwritten

Cancel Import Project

图 30: 导入序列表, 选择常规信息

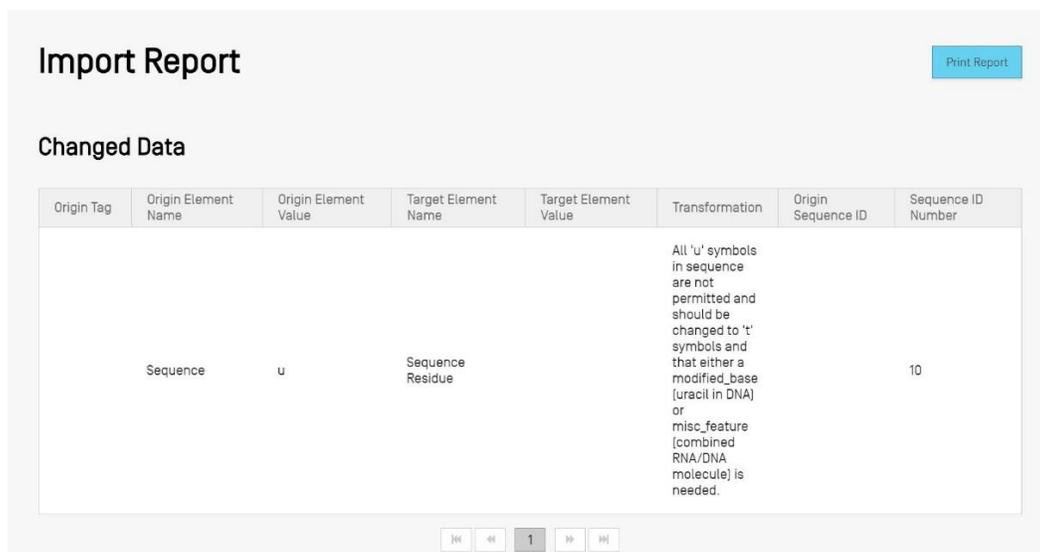
勾选第二个复选框将启用一个附加复选框列表, 这些附加复选框允许用户逐个选择要导入或忽略的属性, 如上图 30: 所示。

3) 最后, 单击蓝色的“导入项目”按钮, 创建新项目。

如果序列表导入正确, 将显示图 31 图 3131 所示的“Changed Data”表, 告知用户在导入过程中对 ST.25、多序列和 raw 序列表数据进行的自动更改, 以使其适应 ST.26 的要求。

注: 导入序列表时, 特征和限定符是区分大小写的, 并应符合产权组织 ST. 26 附件一规定的值。

还应注意的是, 导入的符合 ST. 25 的序列表必须是有效的, 否则不能保证 WIPO Sequence 的功能。



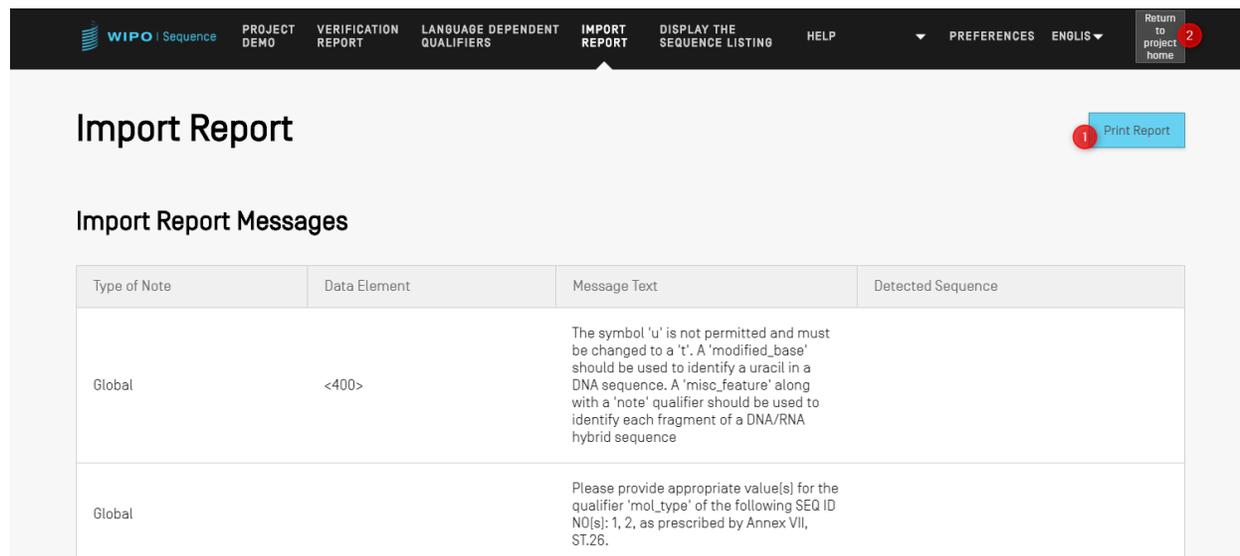
The screenshot shows the 'Import Report' interface. At the top right, there is a 'Print Report' button. Below the title, the section 'Changed Data' contains a table with the following data:

Origin Tag	Origin Element Name	Origin Element Value	Target Element Name	Target Element Value	Transformation	Origin Sequence ID	Sequence ID Number
	Sequence	u	Sequence Residue		All 'u' symbols in sequence are not permitted and should be changed to 't' symbols and that either a 'modified_base' [uracil in DNA] or 'misc_feature' [combined RNA/DNA molecule] is needed.		10

At the bottom of the table, there are navigation controls including a page number '1' and arrows for navigation.

图 31: 导入序列表, 变更数据

如果文件格式是 **ST.25**, 那么导入报告视图将首先包括导入报告表和变更数据表。变更数据报告示例如图 3131 所示, 导入报告示例如图 2532 所示。



The screenshot shows the 'Import Report' interface with a navigation bar at the top. The 'Import Report Messages' section contains a table with the following data:

Type of Note	Data Element	Message Text	Detected Sequence
Global	<400>	The symbol 'u' is not permitted and must be changed to a 't'. A 'modified_base' should be used to identify a uracil in a DNA sequence. A 'misc_feature' along with a 'note' qualifier should be used to identify each fragment of a DNA/RNA hybrid sequence	
Global		Please provide appropriate value(s) for the qualifier 'mol_type' of the following SEQ ID NO[s]: 1, 2, as prescribed by Annex VII, ST.26.	

At the top right of the interface, there is a 'Print Report' button with a red circle containing the number '1' next to it. In the navigation bar, there is a 'Return to project home' button with a red circle containing the number '2' next to it.

图 25: 导入序列表, 导入报告

此时, 用户可以返回项目主页视图 (2) 或以 PDF 格式打印一份关于这些变更的报告 (1): 见图 25。

有关如何下载 PDF 文件的说明, 请参阅图 5。

反过来，如果序列列表文件中存在错误，则导入过程可能失败。在这种情况下，尝试导入后，用户将得到表示导入过程中出现错误的红色横幅的通知，见图 33。

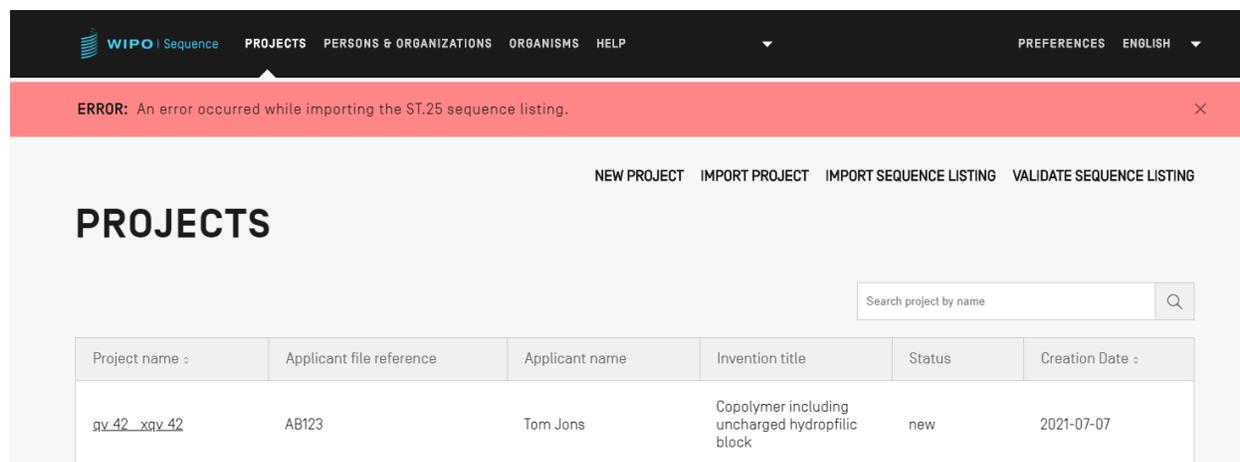


图 33: 导入 ST.25 序列列表，红色横幅

此外，该工具在 100k 序列的阈值限制下表现最佳。当处理大的序列列表时，用户可以采用以下变通方法：通过选择要导入的序列的特定范围，然后将项目中的这些序列按范围导入，从而将导入过程拆分为一系列的步骤。例如，一个约 100k 的序列列表可以被分为 10 x 10k 序列，从而逐一导入。在创建项目时将使用第一个 10K。

验证序列列表

用户可以通过单击项目视图右上方的“VALIDATE SEQUENCE LISTING”按钮来验证 ST.26 序列列表文件，如图 26 所示。

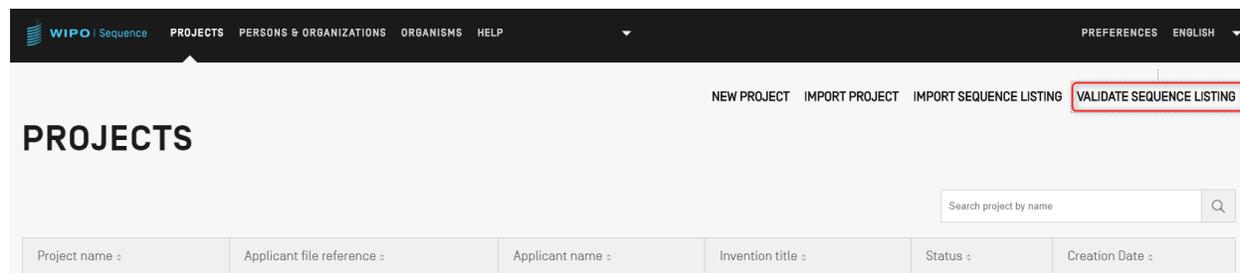


图 26: 验证序列列表

接下来，用户必须如**错误!未找到引用源。**所示，单击“Upload file ST.26 [.xml]”按钮 (1)，然后在对话框中选择文件 (2)。最后，单击“Validate Sequence Listing”按钮 (3)。

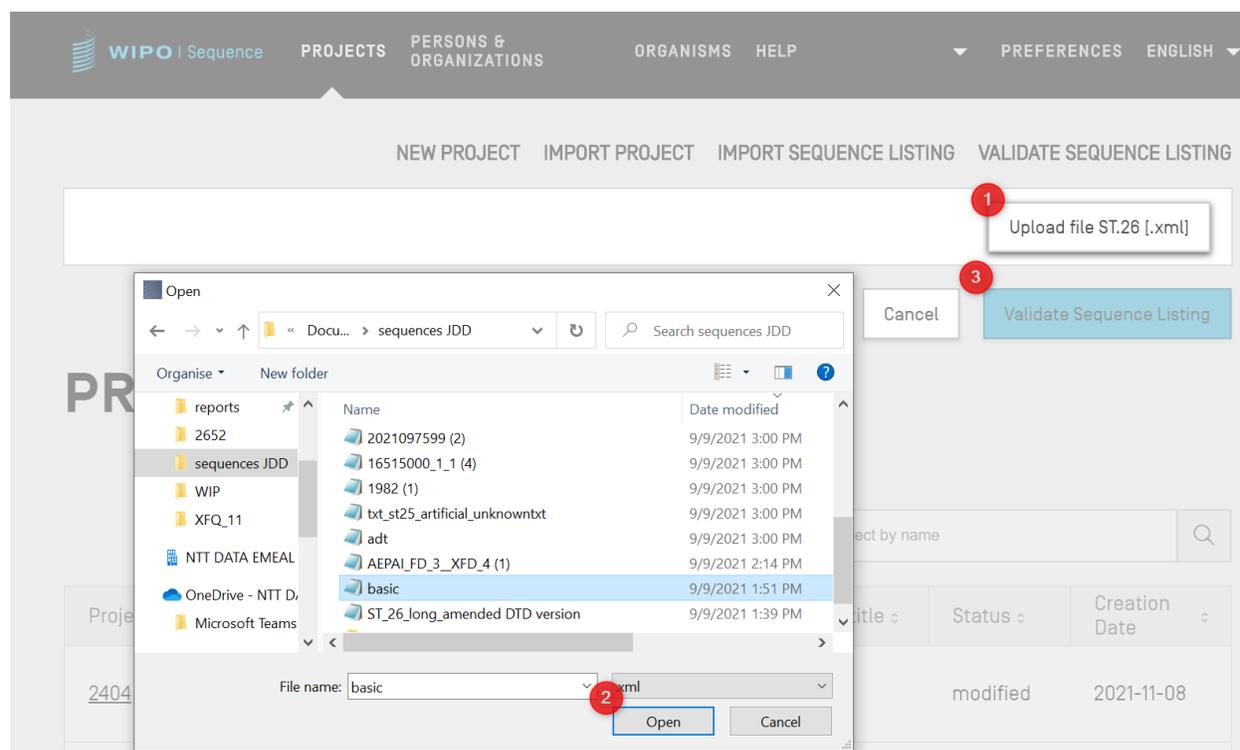


图 27: 验证序列列表, 对话框

如果序列列表通过验证, 就会出现图 286 所示的横幅:



图 286: 横幅显示验证成功完成

如果序列列表未通过验证, 则会在用户浏览器中打开验证报告, 并在表中列出验证错误, 如图 29 所示。

Verification report

Verification Report Information

Production Date	2022-02-12
1 Verification report (XML)	C:\Users\aterrass\AppData\Roaming\ST26_authoring\QT05.xml
2 Verification report (HTML)	C:\Users\aterrass\AppData\Roaming\ST26_authoring\QT05.html

Verification Messages

Severity	Data Element	Message Text	Detected Value	Detected Sequence
ERROR	Qualifier Value	The mandatory qualifier value for qualifier 'note' is missing.		1
WARN	Qualifier Value	The English language value for qualifier 'note' is missing. Certain IP Offices require English language qualifier values.		1
ERROR	Qualifier Value	The mandatory qualifier value for qualifier 'note' is missing.		1
WARN	Qualifier Value	The English language value for qualifier 'note' is missing. Certain IP Offices require English language qualifier values.		1
ERROR	Qualifier Value	The element includes non-permitted characters: м,о,л,е,к,у,а. Only printable characters (including the space character) from the Unicode Basic Latin code table (except the reserved characters) are permitted.	молекула	1
ERROR	Feature Key	The feature key SOURCE is not valid for amino acid sequences. Feature keys for amino acid sequences must be selected from WIPO ST.26 Annex I, Section 7.	SOURCE	2
ERROR	Feature Qualifiers	Mandatory qualifier mol_type is missing.	MOL_TYPE,ORGANISM	2
ERROR	Feature Qualifiers	Mandatory qualifier organism is missing.	MOL_TYPE,ORGANISM	2
ERROR	Qualifier Name	The qualifier name MOL_TYPE is not valid for this SOURCE feature.	MOL_TYPE	2
ERROR	Qualifier ID	This qualifier contains an id attribute. The qualifier id attribute is permitted only for a qualifier with a language-dependent free text value.	q5	2
ERROR	Qualifier Name	The qualifier name ORGANISM is not valid for this SOURCE feature.	ORGANISM	2
ERROR	Non English Qualifier Value	Non-English qualifier free text is permitted only for a qualifier that allows language-dependent free text.	Человек	2

图 29: 验证序列表, 验证报告

HTML 文件的位置将显示在 XML 验证报告 (1) 和 (2) 旁, 以使用户将文件复制到其他位置 (见错误!未找到引用源。)。

注: 用户必须允许内部脚本在其机器上运行, 才能正确显示格式: IE 浏览器中的 ActiveX 控件。这必须设为: “allow blocked content (允许被阻止的内容)”, 以便格式能够正确加载。否则序列将不会以标准格式显示, 而且可读性降低。

请注意, 为了验证序列表, ST.26 文件应符合以下要求:

- 必须使用 UTF-8 编码, 并且必须包含符合 XML 1.0 规范的有效字符
- 必须包含以下 DOCTYPE 行:

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

- 属性 dtdVersion 应符合 DTD 版本:
- 文件必须符合 DTD 文件 ST26SequenceListing_V1_3.dtd⁴。

⁴这是本文件出版时的 DTD 当前版本。

删除项目

要删除项目，用户必须从项目（主页）视图开始，如图 308 所示。

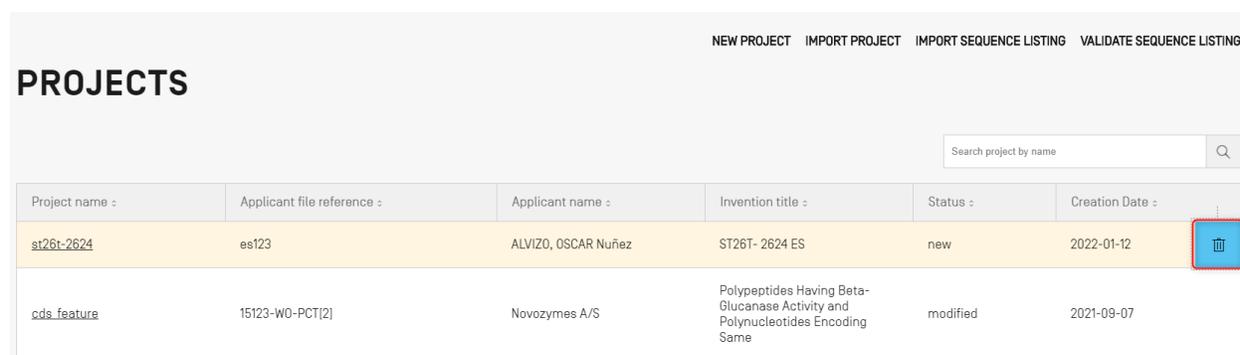


图 30：删除项目，项目主页视图

在项目主页视图表中单击用户希望删除的行上带有回收站图标的按钮。

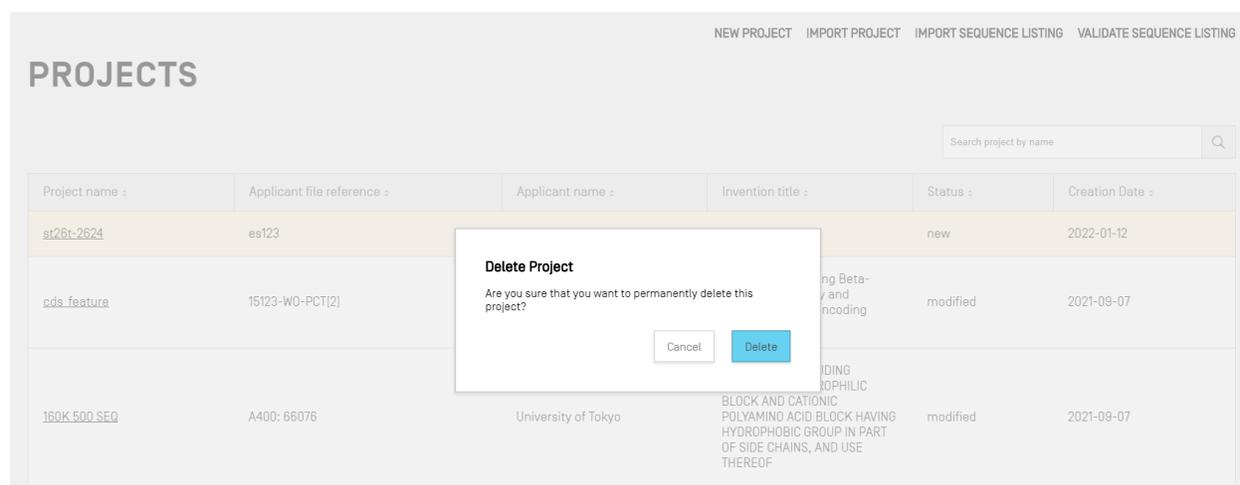


图 31：删除项目，项目主页视图

在**错误!未找到引用源。**所示的弹窗中，单击“Delete（删除）”，确认要删除所选项目。

人员和组织

本节详细介绍人员或组织视图中提供的各种功能。

创建人员或组织

要创建新的人员或组织，用户必须从人员或组织视图开始。

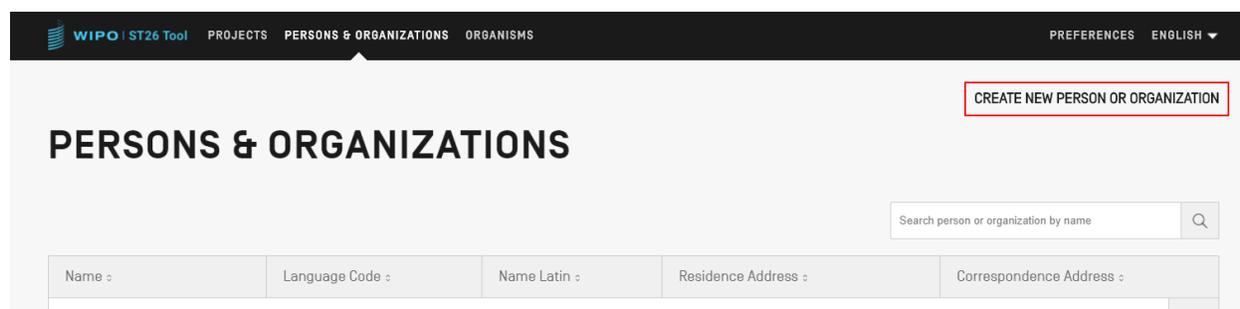


图 32: 创建新的人员或组织第 1 步

首先，用户必须单击视图顶部的“CREATE NEW PERSON OR ORGANIZATION”链接，如图 33 所示。

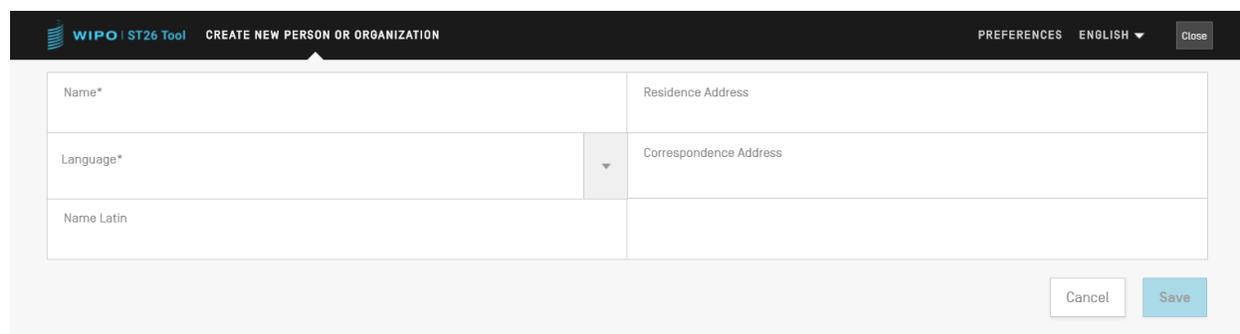


图 33: 创建新的人员/组织第 2 步

在新的视图中，用户必须至少填写与新的人员/组织的详细信息相对应的必填字段（用“*”表示）。对于申请人/发明人而言，只需填写名称（如果以拉丁字符提供）和语言。

如果人员或组织的名称不是拉丁字符，则应在“Name Latin”字段中提供该名称的拉丁文版本。如果未提供此信息，那么在验证或生成 ST.26 序列列表时，项目将无法验证。

自定义生物体

要创建、编辑、导入、导出或删除自定义生物体，用户必须从生物体视图开始。关于如何创建、导入或导出自定义生物体的详细信息在下[错误!未找到引用源。](#)中提供。

创建自定义生物体

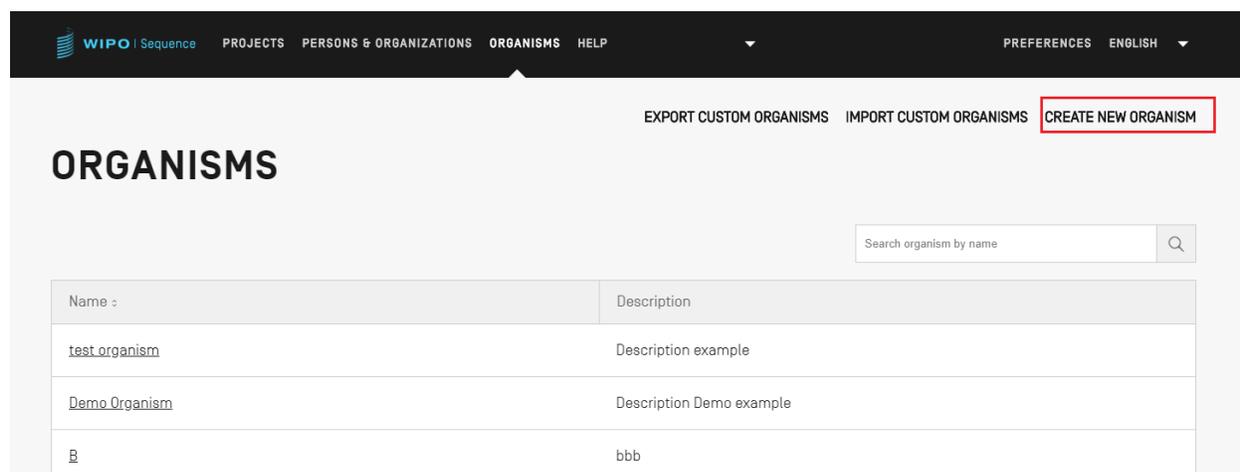


图 34: 创建新的生物体第 1 步

要创建新的自定义生物体，请单击视图顶部的“CREATE NEW ORGANISM”链接，如图 34 所示。在接下来的界面中（图 35），输入新生物体的名称，然后单击“Save”。如果需要对该自定义生物体进行说明，可以进行选择性地添加，如错误!未找到引用源。所示。

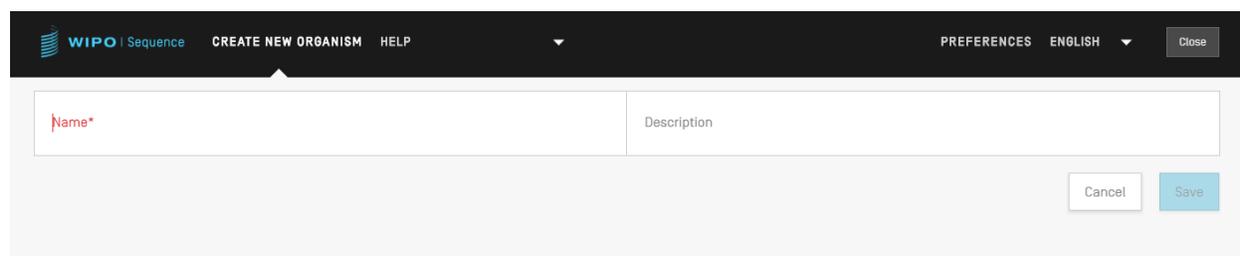


图 35: 创建新的生物体第 2 步

导出自定义生物体

存储在该工具中的所有自定义生物体及其说明都可以导出并保存到文本文件中，以供在该工具之外进行修改或供以后导入。要导出该列表，首先选择“EXPORT CUSTOM ORGANISMS”，如图 36 中所示：

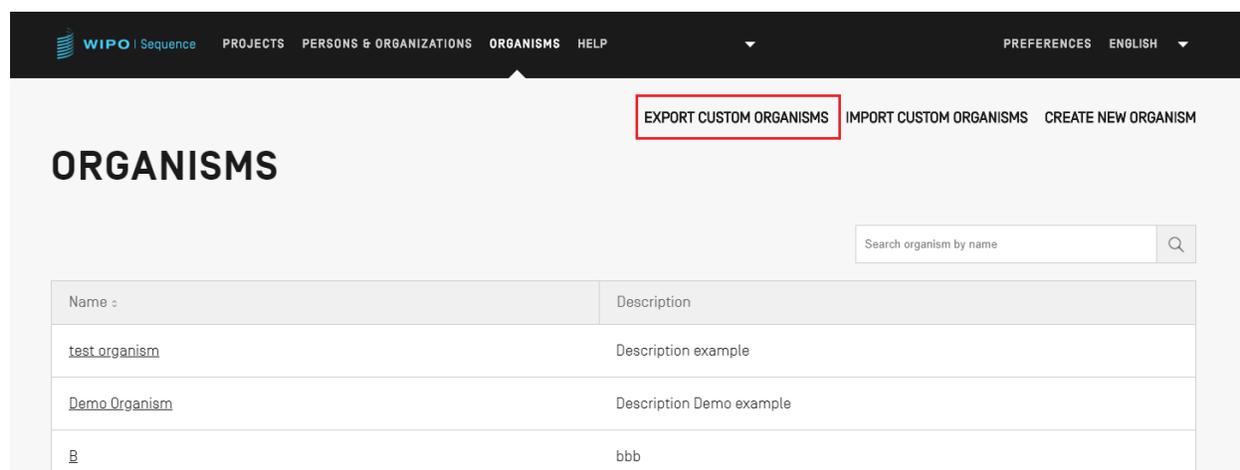


图 36: 导出自定义生物体第 1 步

接下来，将打开一个对话框，允许用户选择文件名和想选的文件位置，如图 37 所示。

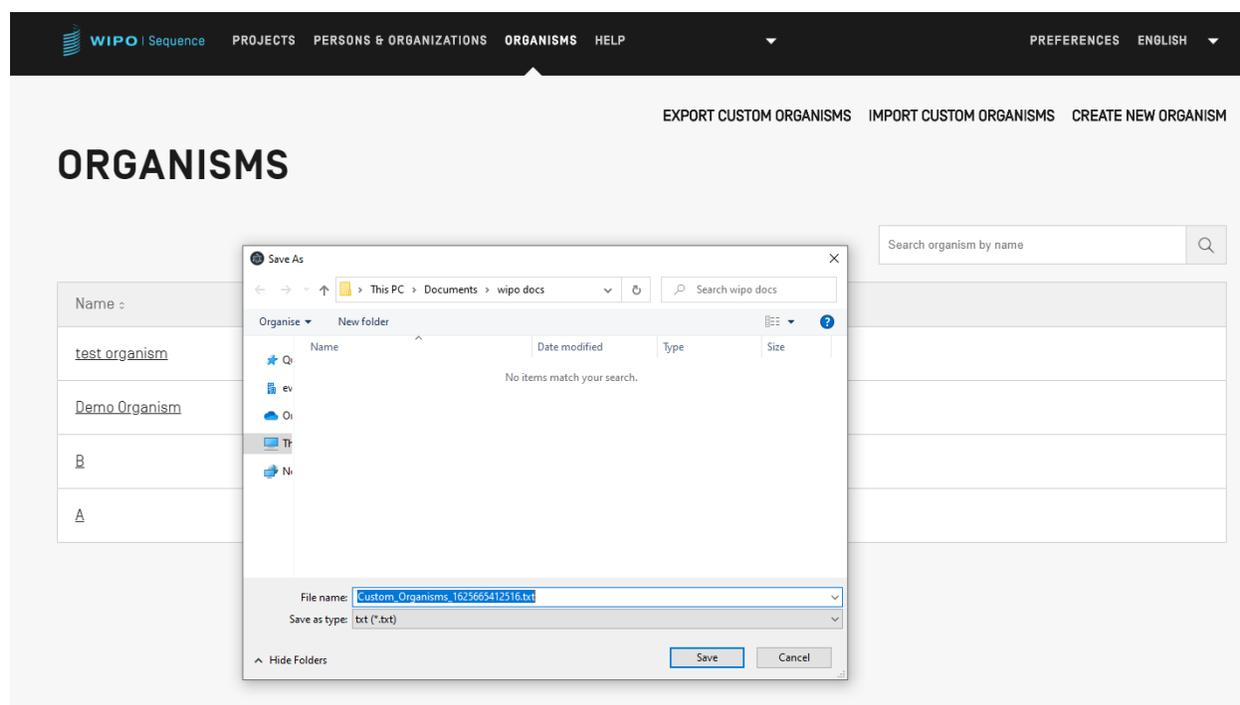
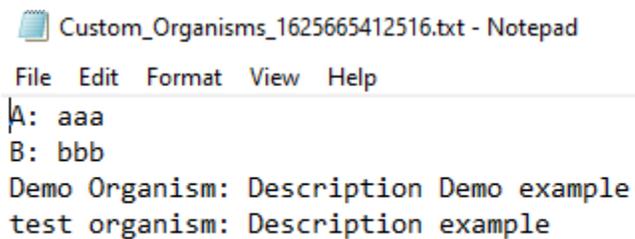


图 37: 导出自定义生物体，对话

导出的文件是 txt 文件，包括生物体的名称和说明，可以编辑并导入工具中（见图 38）。



```
Custom_Organisms_1625665412516.txt - Notepad
File Edit Format View Help
A: aaa
B: bbb
Demo Organism: Description Demo example
test organism: Description example
```

图 38: 自定义生物体文本文件样例

导入自定义生物体

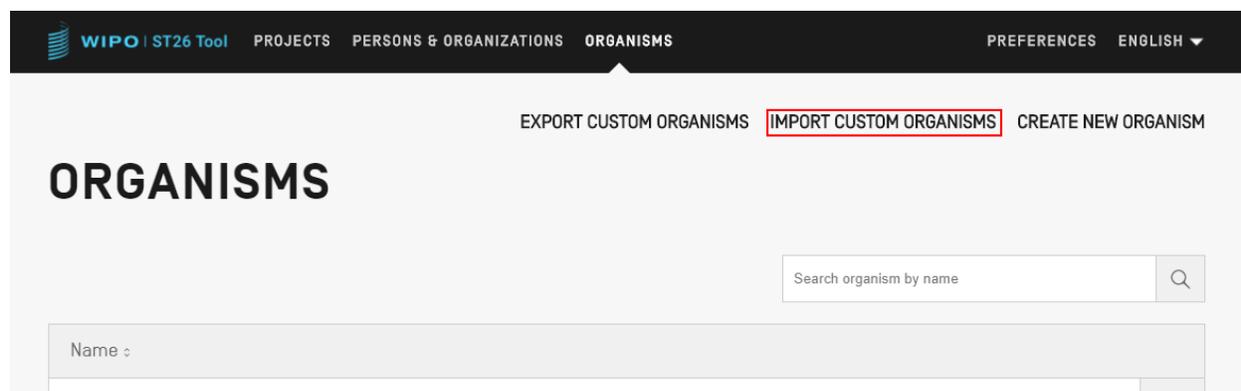


图 39: 导入自定义生物体第 1 步

首先，为了导入自定义生物体列表，用户必须单击视图顶部的“IMPORT CUSTOM ORGANISMS”链接，如图 39 所示。这将在生物体表下方打开一个叠加层，如图 40 所示。

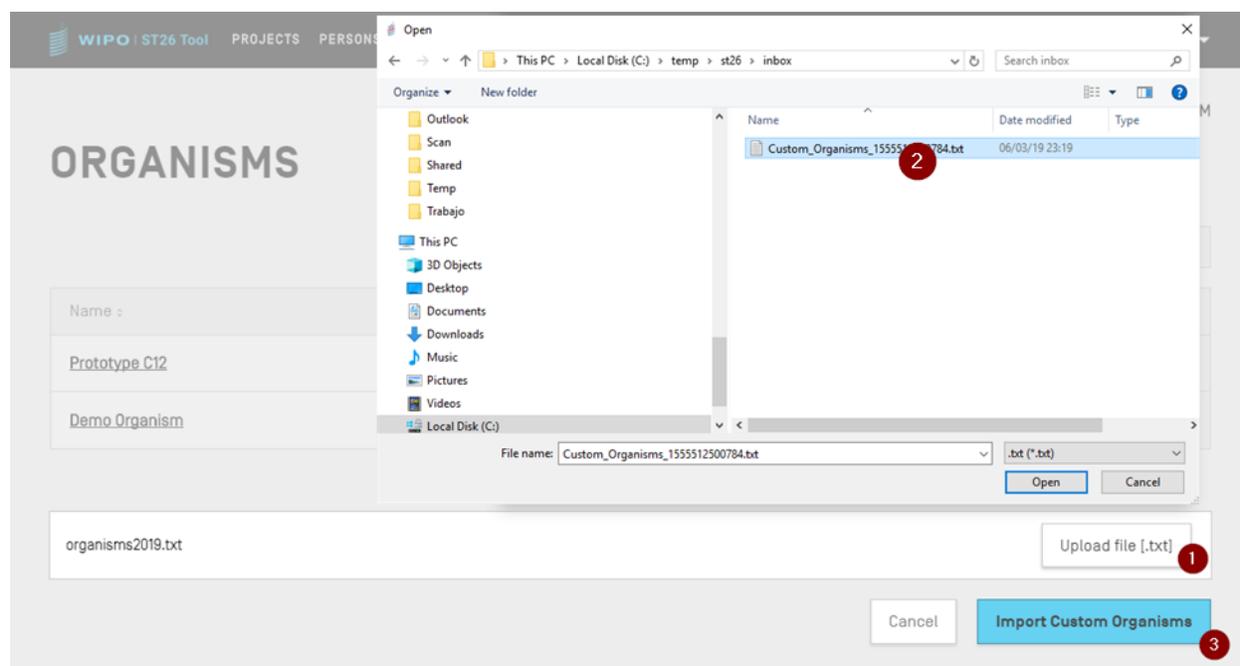


图 40：导入自定义生物体，对话

- 1) 单击“Upload file [.txt]”按钮 (1)
- 2) 从对话框中选择带有自定义生物体名称的文件 (2)
- 3) 最后，单击蓝色的“Import Custom Organisms”按钮 (3)

注：

要导入的文件将是文本文件 (*.txt)，其中包含纯文本 (UTF-8) 形式的自定义生物体名称表，每个项目在一个新行上。

系统首选项

在系统首选项视图中，可以修改 WIPO Sequence 的几个配置参数，这些参数将适用于该工具创建或编辑的每个项目（见图 41）。

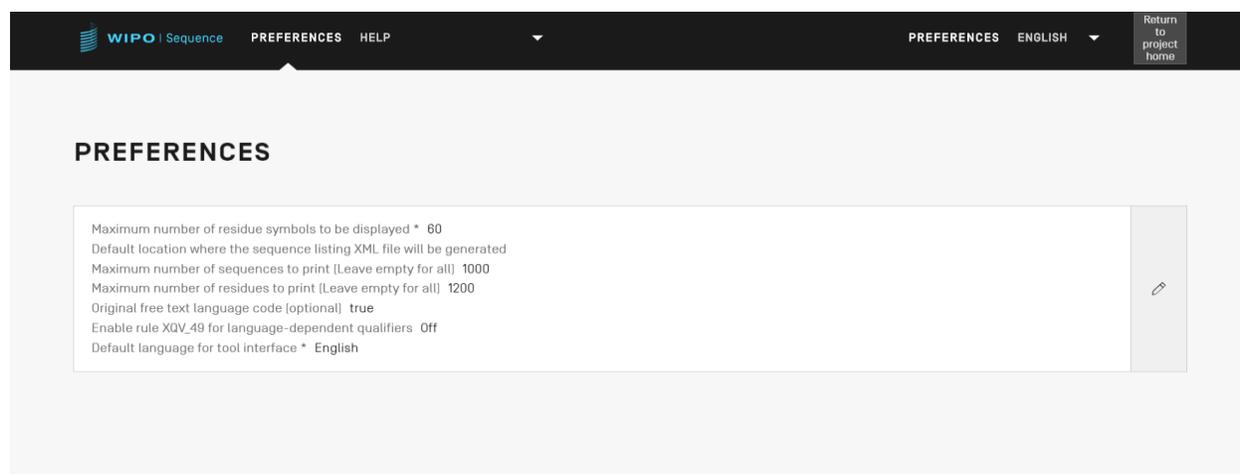


图 41: 系统首选项摘要

为了修改系统首选项，用户应单击上面所示的铅笔图标，以打开图 42 所示的编辑面板。

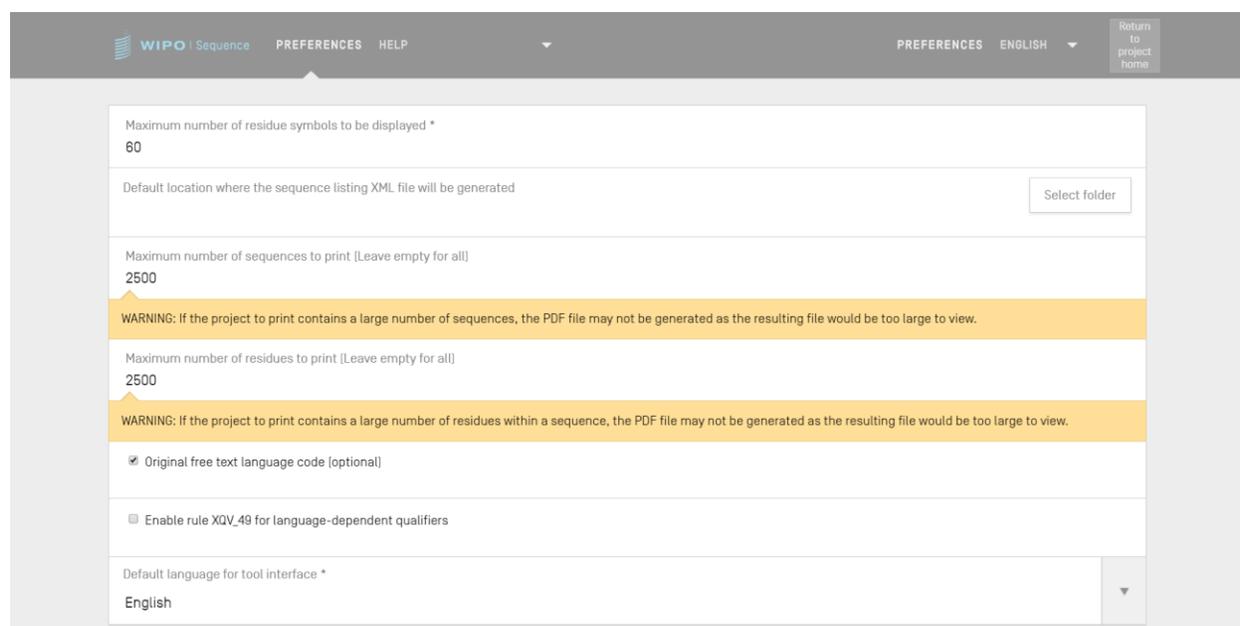


图 42: 配置系统首选项

可以从这个视图中修改的配置项目（按顺序）有：

- **待显示的残基符号数目上限：** 该参数对显示序列时每行显示的残基数目进行设置。默认为 60 个残基；
- **生成 ST.26 序列文件 (.xml) 的默认位置：** 不需要提供该位置。
- **待打印的序列数目上限（留空以选择全部）：** 默认为 1000 个序列。

- **待打印的残基数目上限（留空以选择全部）：**默认为 1200 个残基。
- **原始自由文本语言代码：**如果勾选该复选框，若没有提供原始自由文本语言代码，则验证期间将发出警告。默认情况下未勾选。
- **启用 XQV_49：**如果勾选该复选框，若没有提供语种相关自由文本限定符的英文值，则发出警告。默认情况下未勾选。
- **默认界面语言：**启用 WIPO Sequence 时界面将显示的语言。默认为英文。

注：

第 3 项和第 4 项与将项目打印为 PDF 有关。用户应注意，对于非常大的序列表，产生的 PDF 可能有几千页，无法显示。

项目详细信息

本节介绍项目详细信息视图提供的各种功能。

打印项目

要打印项目，用户必须进入所需项目的项目详细信息视图，并单击视图顶部的“Print”按钮，如错误!未找到引用源。所示。

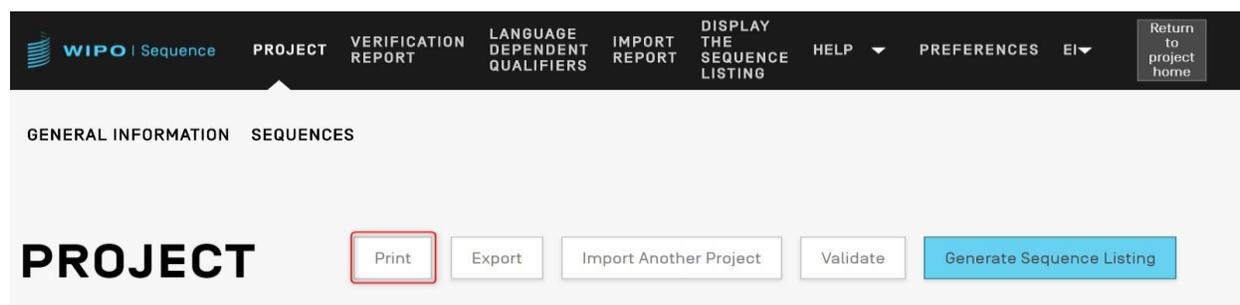


图 43: 打印项目第 1 步

接下来，将向用户显示两个复选框，以澄清用户希望从项目中打印哪些信息：常规信息和/或序列信息（参见图 44）。

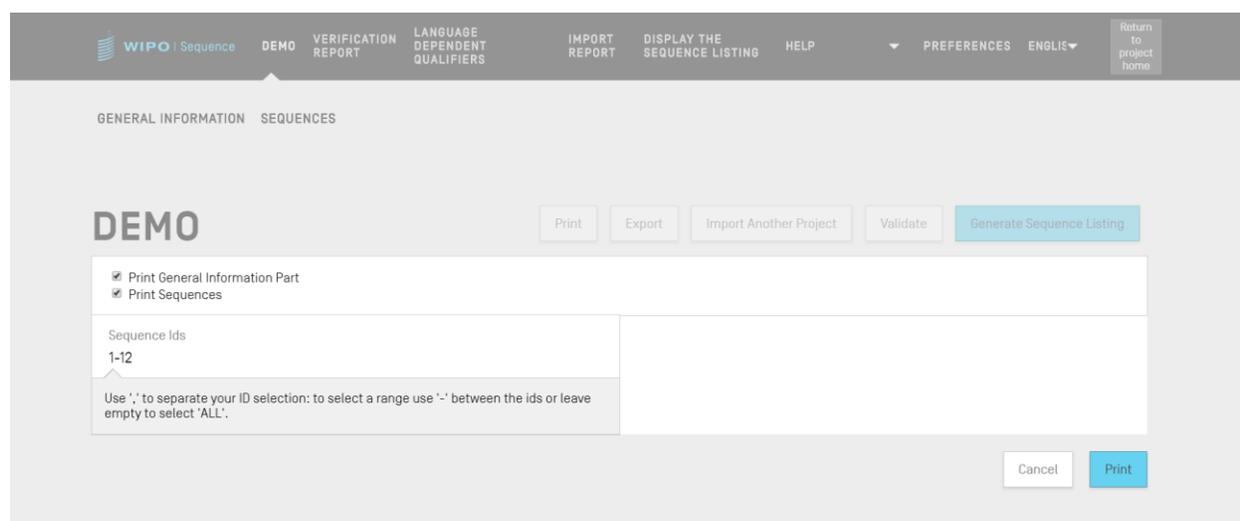


图 44: 打印项目第 2 步

如果选择了“Print Sequences”，用户可以通过在“Sequence IDs”字段中指定 ID 号范围来指定要打印的序列，或者，如果该字段留空，则直接打印所有序列。

默认情况下，项目的序列总数将显示为一个范围。

单击蓝色的“Print”按钮后，如果 PDF 文件生成正确，该工具将在 PDF 阅读器中打开文件，供用户查看。

有关如何下载 PDF 文件的说明，请参阅图 9。

导出项目

项目可以导出为.zip 文件，供用户备份项目数据，或者用另一台安装了 WIPO Sequence 的台式计算机导入，根据图 45 所述的步骤进行。

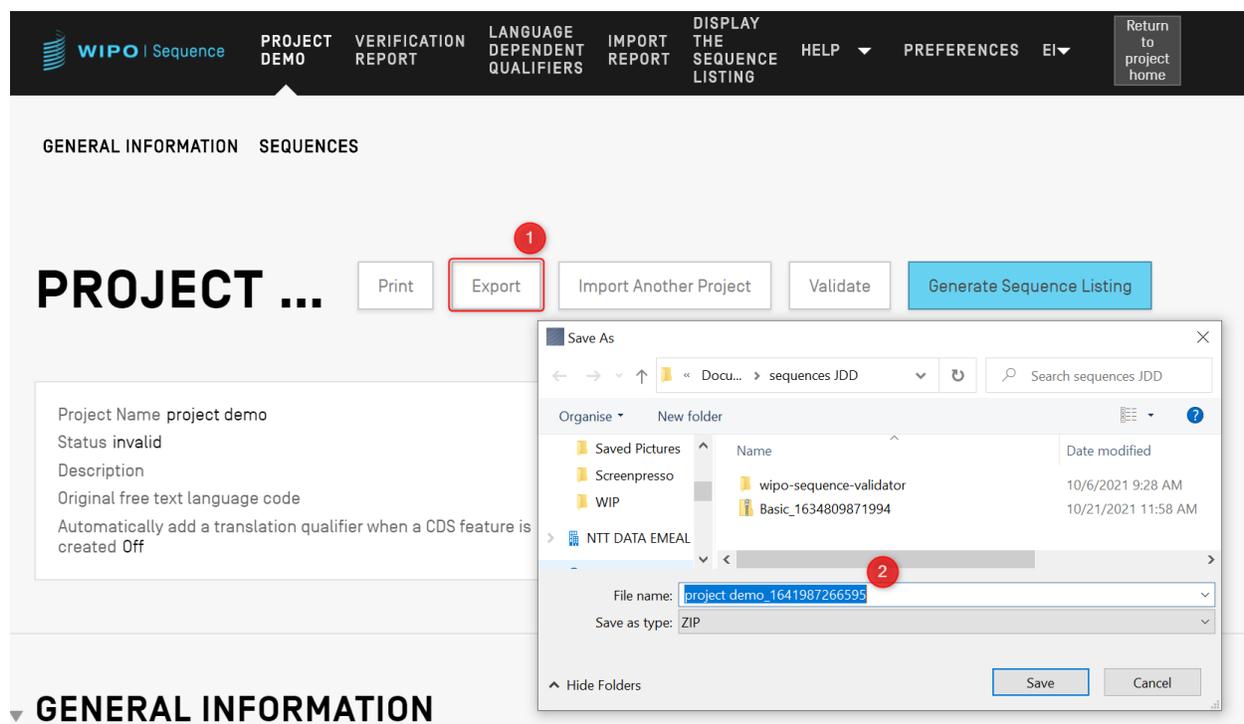


图 45: 导出项目，对话框

- 1) 单击项目详细信息视图顶部的“Export”按钮。
- 2) 在出现的对话框中，选择文件名和所需的项目保存位置。

如果项目成功导出，屏幕上方会出现一个蓝色的横幅，上面有保存的名称和位置，如图 46 所示。

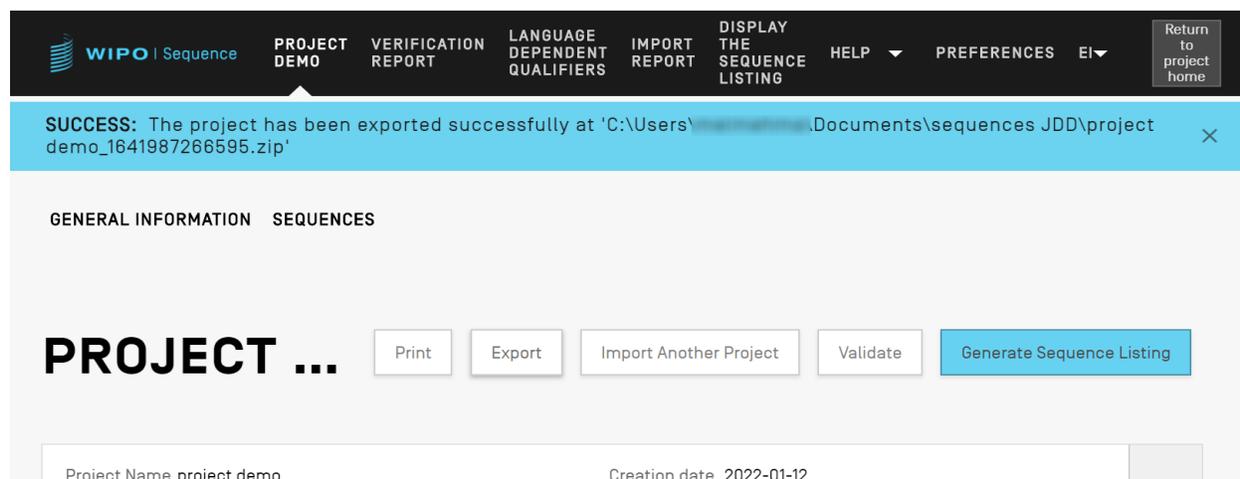


图 46: 导出项目, 成功

从另一个项目导入信息

用户可以将存储在工具中的其他项目的信息复制到当前打开的项目中。此导入的信息可以是“General Information”节，可以是“Sequences”节，或者两者都有。

注：导入的“General Informaiton”将取代项目中当前存在的“General Informaiton”，而导入的“sequences”将附加到项目中当前的序列表中。

首先，用户必须单击项目详细信息视图顶部的“Import Another Project”按钮，如**错误!未找到引用源。**所示。

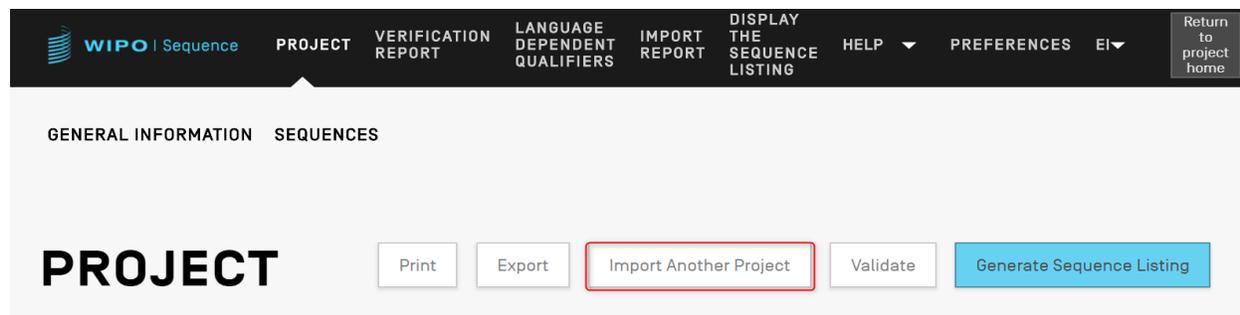


图 47: 导入另一个项目第 1 步

该工具将打开一个叠加层，如**错误!未找到引用源。** 6 所示。用户必须首先选择希望从中导入信息的项目。

BASIC

Project*

Select the general information contents to be imported

Select the sequences to be imported. If the box is left unchecked then no sequences will be imported.

Print Export Import Another Project Validate Generate Sequence Listing

Cancel Import Project

图 48: 导入另一个项目第 2 步

接下来，如**错误!未找到引用源。**所示，用户可以选择是否希望包括项目的“General Information”（1）节中提供的某些详细信息，以及如果希望导入序列（2），可通过输入序列 ID 号（3）范围，以指明将哪些序列导入到项目中。

默认情况下，项目的序列表总数显示为一个范围。

PROJECT DEMO

Project*

1941-AMBIGUOUS

1 Select the general information contents to be imported

2 Select the sequences to be imported. If the box is left unchecked then no sequences will be imported.

Select Range of Sequence IDs.	Total Sequences
1-6	6

Use commas to separate individual sequences [for example: "2,5,8"]; use a dash to indicate a range of sequences [for example: "2-8"].

Print Export Import Another Project Validate Generate Sequence Listing

Cancel Import Project

图 49: 导入另一个项目第 3 步

如果勾选常规信息复选框，会出现一个表格，显示两个项目——当前选中的（原始）项目和目标项目（目的）——的所有常规信息部分。

注：

如果用户没有勾选“Select the general information contents to be imported”框，则不会将序列表常规信息导入到项目中。

然后，用户必须选择哪些“General Information”元素要被相应的目标项目的常规信息所取代，如**错误!未找到引用源。**所示。

<input type="checkbox"/>	Element	Origin Element Value	Target Element Value
<input checked="" type="checkbox"/>	Application Identification	IP Office = FR Application number = 123123123 Filing date = 2022-04-18	IP Office = ES Application number = 1 Filing date = 2021-03-01
<input checked="" type="checkbox"/>	Applicant File Reference	Applicant file reference = 123124	Applicant file reference = ABC1234
<input checked="" type="checkbox"/>	Earliest Priority Application Identification	IP Office = FR Application number = 123123 Filing date = 2022-04-18	IP Office = ES Application number = 1 Filing date = 2021-03-01
<input checked="" type="checkbox"/>	Applicant Name	Name = AP-HM Language code = fr Name Latin = AP-HM	Name = Steven Language code = es
<input checked="" type="checkbox"/>	Inventor Name		Name = Steven Language code = es
<input checked="" type="checkbox"/>	Applicant List	Applicant name = AP-HM	Applicant name = Steven
<input checked="" type="checkbox"/>	Invention Title Bag	Invention title = TEST, Language code = fr	Invention title = AMBIGUOUS, Language code = es

The data for the selected attributes will be overwritten

图 50 导入另一个项目第 4 步

最后，当用户已决定要将哪些常规信息元素和序列导入到项目后，必须单击**错误!未找到引用源。**底部所示的蓝色“Import Project”按钮。

SUCCESS: The project has been imported successfully.

图 51：导入另一个项目，成功

如果这些元素被正确导入，就会出现蓝色的横幅，如**错误!未找到引用源。**所示。

验证项目

在将序列表生成为符合 ST.26 的 XML 文件之前，将事先对项目进行验证检查。该步骤通常在生成序列表之前进行，但也可以单独执行。

要验证项目，用户必须单击项目详细信息视图顶部的“Validate”按钮，如**错误!未找到引用源。**所示。

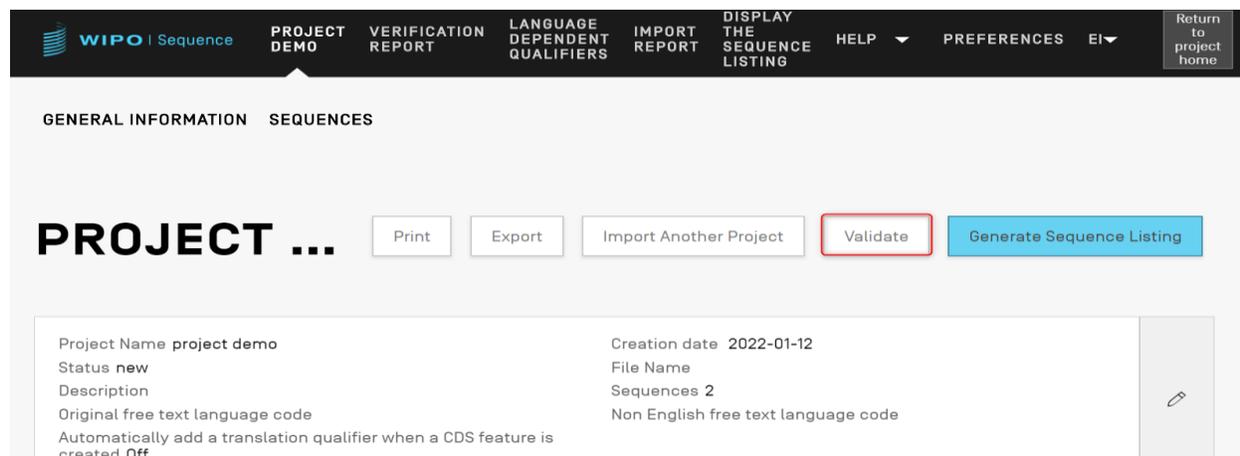


图 52: 项目验证第 1 步

验证完成后，用户将进入“Verification Report”视图，其中显示可能产生的任何验证错误/警告。错误!未找到引用源。是验证成功时显示的屏幕。

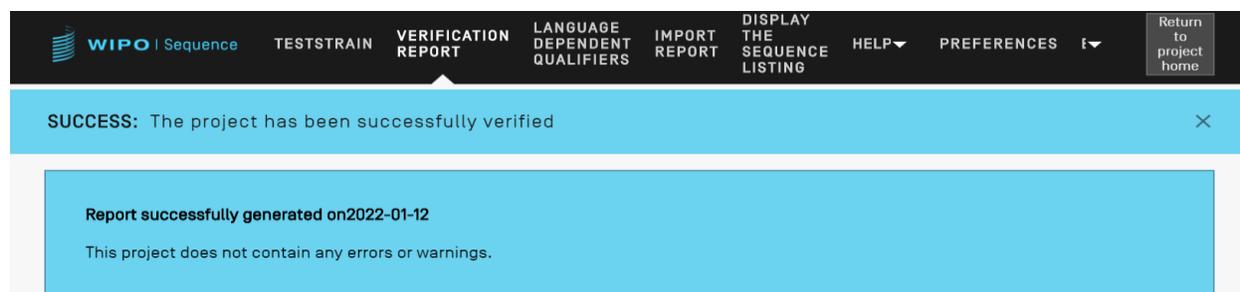
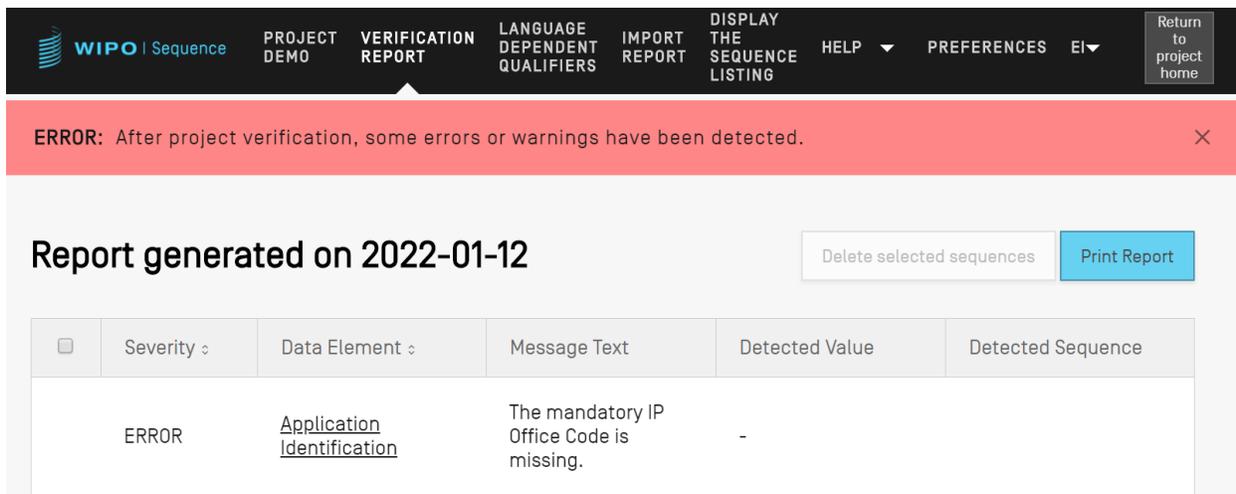


图 53: 项目验证成功

如果验证过程发现任何错误或警告，将生成一份验证报告，其中有一个表格，详细说明了检测到的被违反的验证规则和指导原则。报告示例如错误!未找到引用源。所示，每一行标明了是必须处理的错误，还是用户可以忽略的警告。



The screenshot shows the WIPO Sequence application interface. At the top, there is a navigation bar with the logo and several menu items: PROJECT DEMO, VERIFICATION REPORT (highlighted), LANGUAGE DEPENDENT QUALIFIERS, IMPORT REPORT, DISPLAY THE SEQUENCE LISTING, HELP, PREFERENCES, and a 'Return to project home' button. Below the navigation bar, a red error message banner reads: 'ERROR: After project verification, some errors or warnings have been detected.' Below this, a report header indicates 'Report generated on 2022-01-12' with buttons for 'Delete selected sequences' and 'Print Report'. A table displays the error details:

<input type="checkbox"/>	Severity ◊	Data Element ◊	Message Text	Detected Value	Detected Sequence
	ERROR	Application Identification	The mandatory IP Office Code is missing.	-	

图 54: 项目验证, 错误/警告

生成序列列表

对项目可以执行的最后操作（也许是最重要的操作），是生成序列列表。要生成序列列表，用户必须单击项目详细信息视图顶部的蓝色“Generate Sequence Listing”按钮，如图 55 所示。这将自动触发首先对项目进行的验证过程。

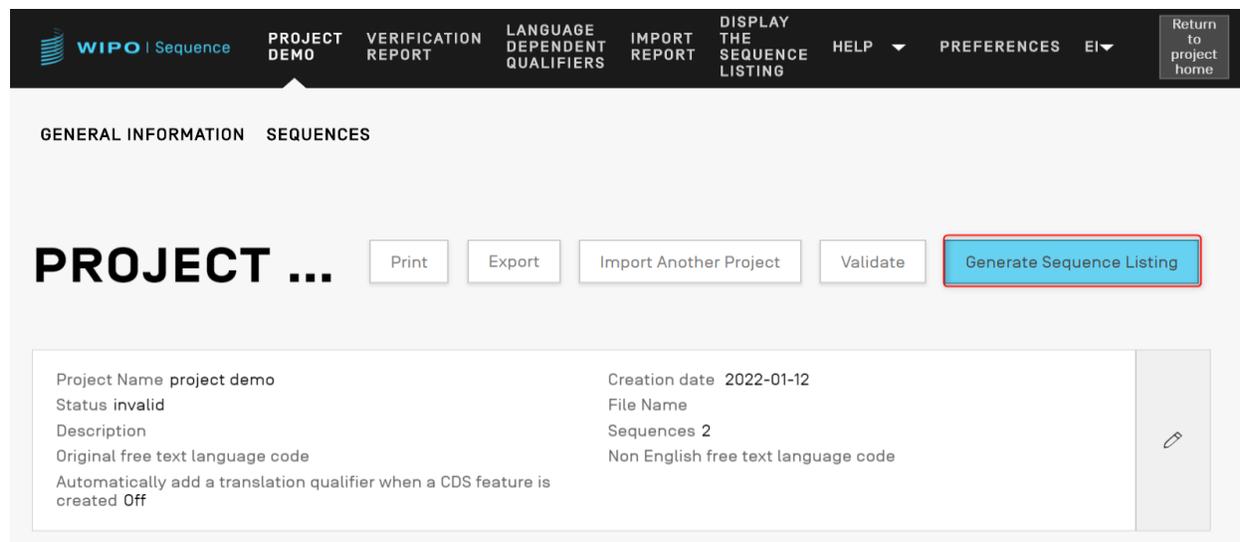


图 55: 生成序列列表

如果项目通过验证过程，将打开一个对话框，让用户选择一个位置来保存生成的符合 ST.26 的序列列表（.xml），如图 56 中显示。

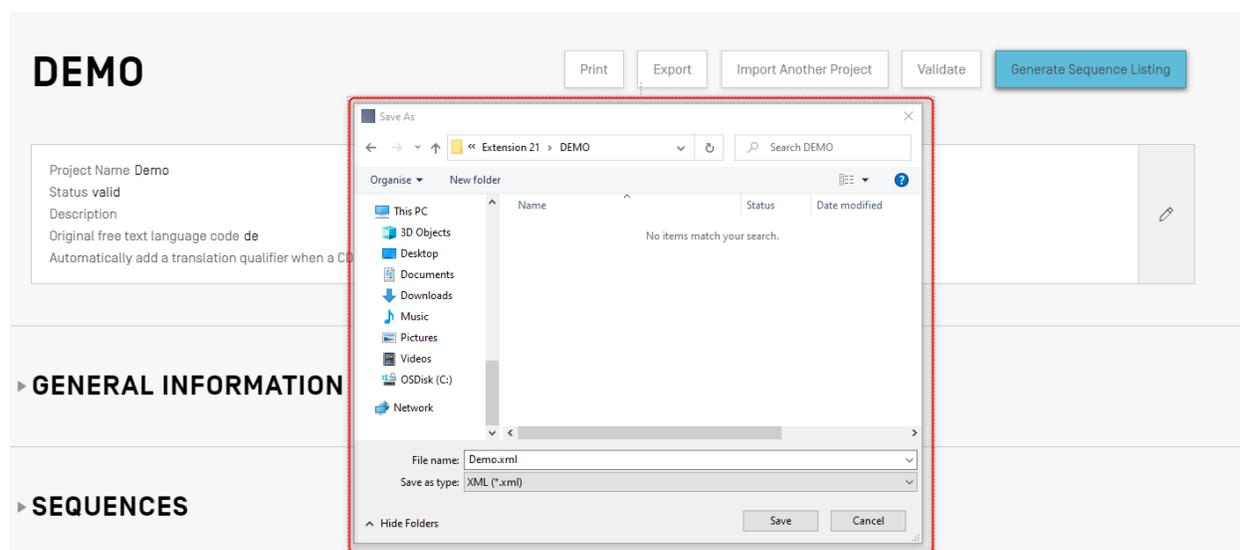


图 56: 生成序列列表, 对话

注:

使用 Linux 发行版时存在一个已知问题: 默认情况下, 建议文件名会出现一个多余的“\”。为解决该问题, 请在保存前手动删除多余的“\”。

如果项目未能通过验证, 那么验证报告视图将改以红色横幅显示, 如图 57 所示。

ERROR: The ST.26 sequence listing file cannot be generated because the project contains errors. ✕

Report generated on 2022-01-21

Delete selected sequences Print Report

<input type="checkbox"/>	Severity ◯	Data Element ◯	Message Text	Detected Value	Detected Sequence
<input type="checkbox"/>	ERROR	<u>Qualifier Molecule Type</u>	The value of the qualifier 'mol_type' is not one of the permitted values for a DNA sequence or the value is missing.	rRNA	Sequence 1
<input type="checkbox"/>	ERROR	<u>Feature Location</u>	The feature location includes a residue number greater than the length of the sequence, which is invalid.	1..50	Sequence 1

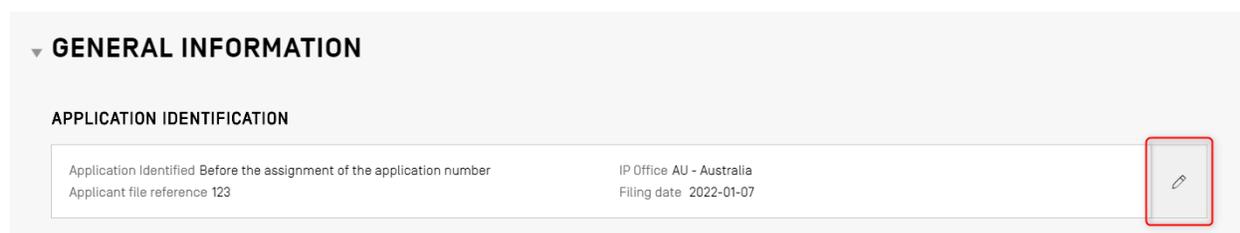
图 57: 生成序列列表, 错误/警告

常规信息

本节允许用户输入与专利申请本身相关的信息，用于将生成的序列表与该申请关联起来。

申请识别

常规信息节的“Application Identification”子节与所选项目的专利申请状态和信息有关，参见图 58。



APPLICATION IDENTIFICATION	
Application Identified Before the assignment of the application number	IP Office AU - Australia
Applicant file reference 123	Filing date 2022-01-07

图 58: 编辑申请识别

要编辑“Application Identification”子节中的信息，单击上方该子节右侧显示的铅笔图标。然后，用户必须按照以下步骤提供信息，如图 59 所示：

- 1) 如果申请已经分配了申请号，用户必须选择其提出申请的知识产权局的代码。这是 WIPO ST.3 代码。
- 2) 用户必须通过选择相应的单选按钮，来选择是否已经收到申请号，或者刚在申请文件中提供。
- 3) 在没有申请号的情况下，用户**必须**在此字段中提供申请人档案号。
- 4) 如果已经分配了申请号，用户应输入为专利提供的申请号。
- 5) 如果已经分配了申请日，请用日期选择器选择申请的提交日期。
- 6) 单击蓝色的“Save”按钮。

GENERAL INFORMATION

APPLICATION IDENTIFICATION

IP Office
BD

Application Identification*
 Before the assignment of the application number
 After the assignment of the application number

Applicant file reference*
123456
Application File Reference is a mandatory field when 'Before the assignment of the application number' is provided.

Application number
12345

Filing date
2019-02-20

Cancel Save

图 59: 申请识别编辑面板

注: 无论输入什么, 验证报告总会出现一条警告, 提示“没有申请识别号。如果已分配申请号, 则申请号必填”。

优先权识别

要在项目中添加优先权申请, 用户必须单击项目详细信息视图中常规信息节的“Add Earliest Priority Identification”按钮, 如图 60 所示。

GENERAL INFORMATION

APPLICATION IDENTIFICATION

Application Identified Before the assignment of the application number
Applicant file reference 123

IP Office AU - Australia
Filing date 2022-01-07

PRIORITY IDENTIFICATION

Add Priority Application

图 60: 编辑/添加优先权申请第 1 步

要将当前选定的优先权申请设置为最早, 用户必须在“Selected Earliest Priority Application”下拉菜单中选择“**Yes**”。这将在生成序列列表时, 将其设置或修改为最早的优先权申请。

要完成此操作, 请单击叠加层中的蓝色“Add Earliest Priority Application”按钮, 如图 61 所示。

PRIORITY IDENTIFICATION	
<button>Add Priority Application</button>	
IP Office*	Filing date Date
Application number*	Selected Earliest Priority Application*

Cancel Add Priority Application

图 61: 优先权申请编辑面板

申请人和发明人

要将新申请人或新发明人的数据添加到项目中，用户必须在项目详细信息视图的常规信息节单击“Add Inventor”或“Add Applicant”按钮。执行这两个操作的步骤是相同的，因此只提供常规说明，但如果申请人和发明人都要添加在项目中，则这个过程必须重复两次，即便申请人同时也是发明人。

这时将打开一个带有两个单选按钮的叠加层，如图 62 所示。如果选择了“Existing applicant/inventor”，用户可以从一个下拉框中进行选择，该下拉框列出了桌面工具本地实例中当前保存的人员和组织。图 62 显示一个列表，其中有三名现有申请人，包括“John Smith”。

Add Applicant Add Inventor

Select the option*

Existing applicant

New applicant

Select applicant*

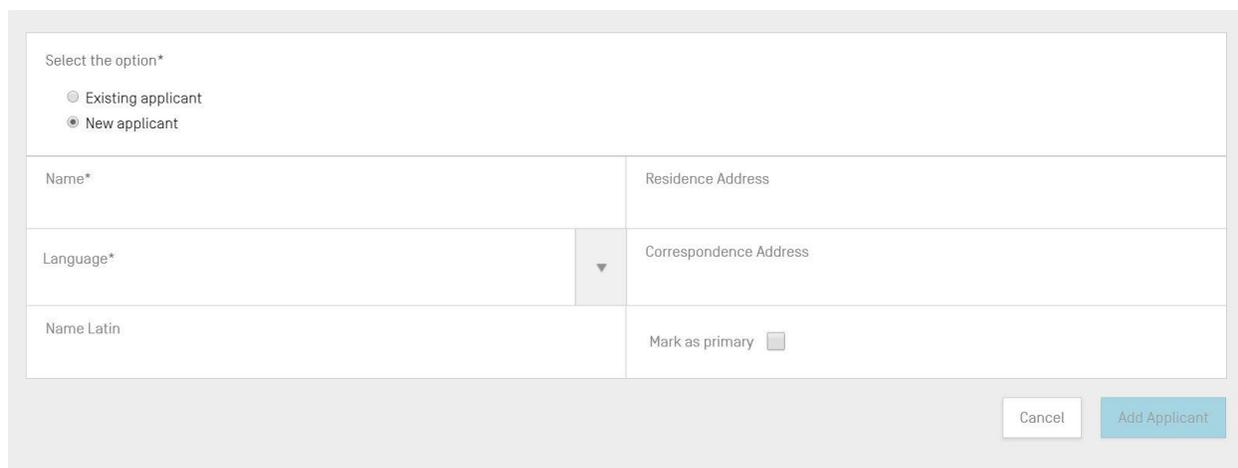
הילה בן אברהם
Juan Rodriguez Garcia
John Smith

图 62: 添加现有申请人/发明人

如果选择了“New applicant/inventor”，用户必须按创建新的人员/组织一样的方式填写编辑面板（参见图 63）。

请：只需一名申请人序列列表即可被视为有效。因此，必须将一名申请人和/或发明人标为 **primary**（主要的）。这是将出现在生成的序列列表中的申请人/发明人。

最后，填完详细信息后，用户应单击图 63 所示的“Add Applicant/Inventor”按钮。由于下面提供的必填字段未填写，蓝色的“Add Applicant”按钮呈灰色。



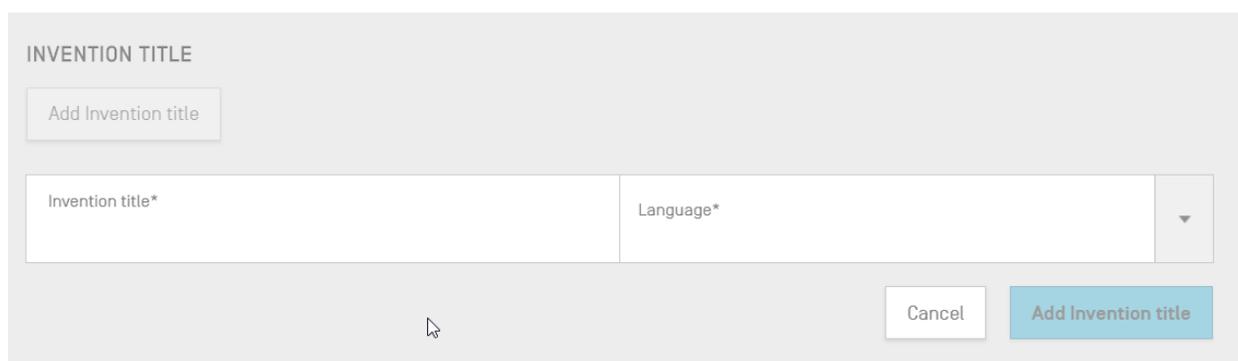
Select the option*	
<input type="radio"/> Existing applicant	
<input checked="" type="radio"/> New applicant	
Name*	Residence Address
Language*	Correspondence Address
Name Latin	Mark as primary <input type="checkbox"/>
<input type="button" value="Cancel"/> <input type="button" value="Add Applicant"/>	

图 63: 添加新申请人/发明人

发明名称

“Invention Title”是常规信息节的最后一个子节。

- 1) 要添加新的发明名称，单击“Add Invention title”按钮，如图 64 所示。



INVENTION TITLE	
<input type="button" value="Add Invention title"/>	
Invention title*	Language*
<input type="button" value="Cancel"/> <input type="button" value="Add Invention title"/>	

图 64: 添加发明名称编辑面板

- 2) 在这个叠加层中，用户必须输入发明的名称，同时还要说明发明名称是用什么语言提供的。

- 3) 单击蓝色的“Add Invention title”按钮。

注：根据 WIPO ST. 26，序列表必须以提交申请的语言提供发明名称。但是，项目也可以选择用其他语言添加多个发明名称，但每种语言只能有一个发明名称。每个新发明名称都可以通过上述步骤添加。

序列

项目详细信息视图的“Sequences”部分是用户提供与序列本身相关的技术信息的地方。

要创建/导入/插入/重排一个序列，用户必须向下滚动到项目详细信息视图底部的“Sequences”节。下面的子节提供了执行这些操作所需步骤的更多信息。

创建序列

- 1) 单击“Create new sequence”按钮，如图 65 所示。

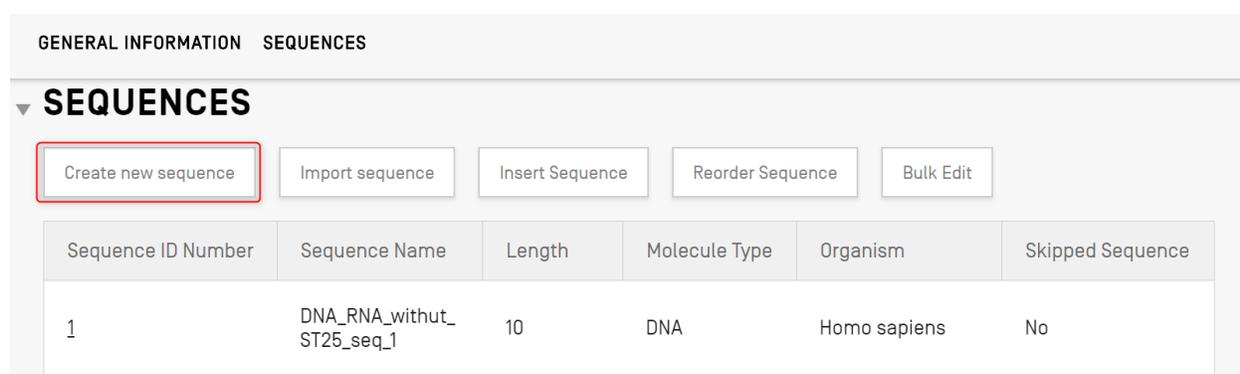


图 65: 创建新序列第 1 步

其余步骤如图 66 所示:

- 2) 用户可以选择通过在“Sequence name”字段中提供值来提供一个名称，以便更容易地区分该序列。如果留空，桌面工具将为每个新序列分配一个默认值，默认值以“Seq”打头，后接迭代递增的数字（“Seq_1” “Seq_2” “Seq_3”）（1）。

用户必须选择 ST.26 允许的并由下拉框提供的三种分子类型之一（‘DNA’ ‘RNA’ 和 ‘AA’）（2）。序列本身必须在“Residue”⁵文本字段中输入（3）。注意，如果用户希望创建一个同时包含 DNA 和 RNA 片段的序列，必须选择 DNA 作为主要分子类型。

用户还必须为要创建的序列选择生物体名称，因为这是必填字段。这可以从该桌面工具数据库中预定义的生物体列表中选择，也可以从用户在本机创建并保存的自定义生物体中选择。用户只需开始输入生物体的名称，就会出现一个下拉列表，其中包含这两个来源的选项（4）。

用户还可以通过从下拉框（5）中选择限定符分子类型⁶来对生物体进行分类，该下拉框提供的值根据先前选择的分子类型而变化。

如果勾选“Mark as an intentionally skipped sequence”（6），序列面板将删除对提供必填元素值的限制，并在验证项目和生成序列时，忽略所产生的已保存序列（残基值将被设置为‘000’）。

如果勾选“The sequence contains both DNA & RNA segments”（6），面板将扩展到包括允许用户用“misc_feature”特征描述每个 DNA 和 RNA 片段的字段。每个已定义片段的位置将被存储到序列的一个不同的“misc_feature”特征（7）中，“misc_feature”特征还包括跟在分子类型后的“Further Text”文本字段值中的“note”限定符。用户可以根据需要，单击“Add new ‘misc_feature’ feature”按钮来创建多个此类特征，而且必须对整个序列中的所有片段（包括 DNA 和 RNA 片段）都这样做。

⁵ 参见产权组织标准 ST.26-附件一第 1 节和第 3 节：有效核苷酸符号表和氨基酸符号表。

⁶ 参见产权组织标准 ST.26-附件一第 6 节和第 8 节：核苷酸表和氨基酸表之下分子类型条目的限定符值。

图 66: 创建新序列编辑面板

- 3) 要完成操作，用户可以单击灰色的“Create sequence”按钮或蓝色的“Create & Display Sequence”按钮。

如果用户单击蓝色的“Create & Display Sequence”按钮，在创建序列后，将打开一个可折叠的序列显示，位于项目详细信息视图中的序列节下方，供用户查看各种值。

新创建的序列可以在序列表的最后一个位置找到，并带有下一个可用的序列 ID 号。新创建的序列样例如[错误!未找到引用源。](#)所示。

要查看关于如何重排序列表的详细信息，参见[错误!未找到引用源。](#)和[错误!未找到引用源。](#)。

▼ **SEQUENCE 2**

Sequence Number (ID) 2	Molecule Type DNA
Sequence Name Name/Description for sequence	Organism Saaristoa firma
Length 1792	Contains DNA and RNA fragments

▼ **FEATURES**

Add feature

Feature Key	Location	Qualifiers
misc_feature	7..13	note = RNA
source	1..1792	mol_type = genomic DNA organism = Saaristoa firma

▼ **SEQUENCE**

atctacttca	ttgaatctac	ttcattgaat	ctacttcatt	gaatctactt	cattgaatct	acttcattga	atctacttca	80
ttgaatctac	ttcattgaat	ctacttcatt	gaatctactt	cattgaatct	acttcattga	atctacttca	ttgaatctac	160
ttcattgaat	ctacttcatt	gaatctactt	cattgaatct	acttcattga	atctacttca	ttgaatctac	ttcattgaat	240
ctacttcatt	gaatctactt	cattgaatct	acttcattga	atctacttca	ttgaatctac	ttcattgaat	ctacttcatt	320
gaatctactt	cattgaatct	acttcattga	atctacttca	ttgaatctac	ttcattgaat	ctacttcatt	gaatctactt	400

图 67: 新序列保存在项目中

导入序列

序列也可以直接从文件导入到项目中。可接受的文件格式⁷有 **raw**、**multi-sequence**、**FASTA**、**ST.26** 和 **ST.25**。勾选后，WIPO Sequence 会自动检测文件中使用的格式。

- 1) 单击“Import sequence”按钮，如图 68 所示。

⁷ 参见本文件的文件格式部分（第 5 节）。

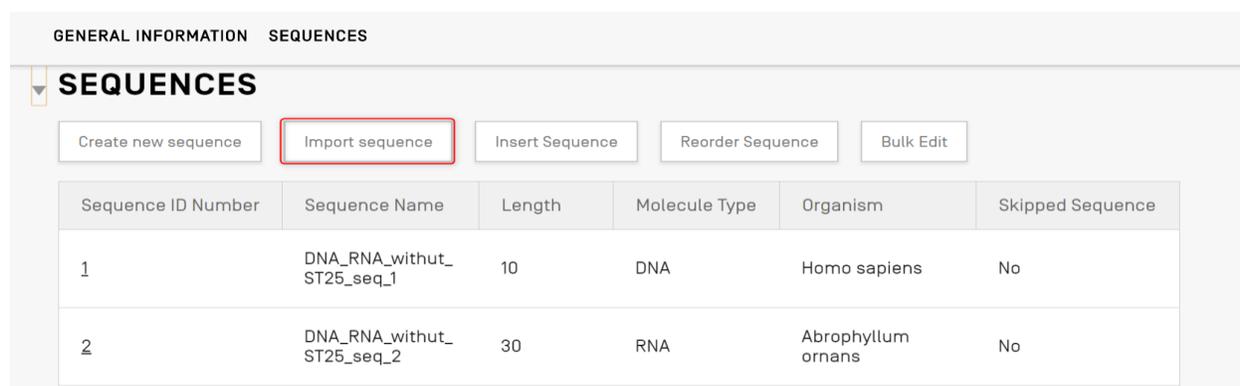


图 68: 导入序列第 1 步

2) 单击“Upload file [.txt, .xml]”，如图 69 所示。

当对话框打开时，选择含有要导入的序列数据的文件。该桌面工具将检测所使用的格式，并对导入执行一些验证检查。

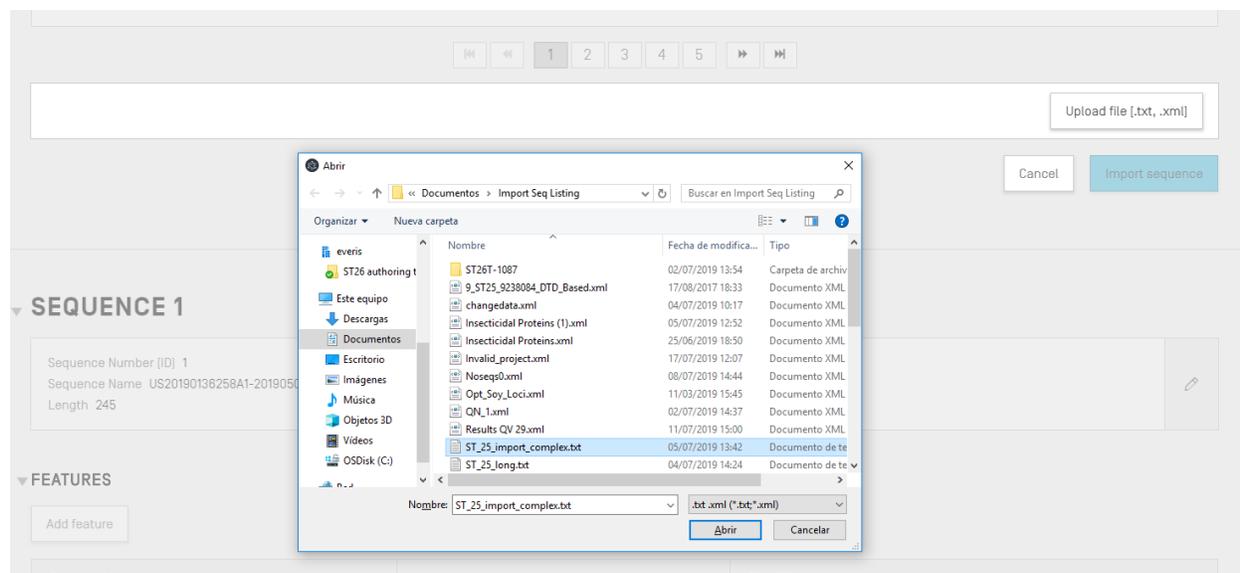


图 69: 导入序列，对话

该工具可接受五种导入序列的格式：**raw**、**multi-sequence**、**FASTA**、**ST.26** 和 **ST.25**。本文件上一节概述了必须采用这些文件的格式。

如果所选文件的格式是 **ST.25** 或 **ST.26**（见图 70），用户将首先看到“**Select Range Sequences**”复选框（1）。勾选后，将打开一个表（2），其中包含文件中每个序列的序列 ID 号，以及它们将被附加到项目所提供的序列列表中的顺序。

如果用户不希望将所有的序列导入到项目中，可以提供所需的序列 ID 号范围（3）。

可以输入单个序列，也可以输入用逗号分隔的序列列表，或者输入 x-y 形式的序列范围。

示例：“1, 3, 7, 13-20, 30-50”。

C:\Users\...Downloads\valid1.xml Upload file (.txt, .xml, .FASTA)

1 Enter the sequences to be imported

Sequence Number [ID]	Position
[1]	1
[2]	2
[3]	3

3 Select Range of Sequence IDs.
1-3
Use commas to separate individual sequences [for example: "2,5,8"]; use a dash to indicate a range of sequences [for example: "2-8"].

Cancel Import sequence

图 70: 导入序列，ST.25/ST.26 格式

在导入多序列格式文件的情况下（见图 71），用户将看到一个“**Select Range Sequences**”复选框（1），勾选该复选框后，将显示一个预览表，显示文件中相应序列的序列 ID 号，并在“**Details**”栏（2）显示每个序列的详细信息：包括序列名称、分子类型和生物体名称。

用户必须选择希望导入到项目列表中的序列 ID 号范围（3）。默认情况下，所选择的项目列表文件的总序列数显示为一个范围。

C:\Users\...Downloads\multisequence_u_t.txt Upload file [.txt, .xml, .FASTA]

Enter the sequences to be imported

Sequence Number [ID]	Position
[1]	< First Sequence;DNA;Abies alba>
[2]	< Second Sequence;RNA; >
[3]	< Second Sequence;AA; >
[4]	< Third Sequence;DNA;Abies alba>

Select Range of Sequence IDs.
1-4

Use commas to separate individual sequences [for example: "2,5,8"]; use a dash to indicate a range of sequences [for example: "2-8"].

Cancel Import sequences

图 71：导入序列，多序列格式

导入序列过程所接受的最后两种格式是 **raw** 文件格式和 **FASTA** 文件格式（参见文件格式一节）。这些格式每个文件仅定义一个序列。选中要导入的 **raw** 文件后，该工具将显示图 72 所示的编辑面板，当导入 **FASTA** 文件时显示图 73 所示的编辑面板。用户应通过提供必填字段进行操作。

C:\Users\aterrass\everis\ST26 authoring tool - 2020-21 Implementation internal - 2020-21 Maintenance\Extension 21\Testing files\3AAraw.txt Upload file [.txt, .xml, .FASTA]

The file selected is in raw format. Please select the sequence molecule type in order to import the sequence.

Sequence Name	Molecule Type*
Organism name *	Qualifier Molecule Type

Cancel Import sequence

图 72：导入序列，raw 格式

C:\Users\laterrass\everis\ST26 authoring tool - 2020-21 Implementation internal - 2020-21 Maintenance\Extension 21\Testing files\FASTA\1DNAsequence.fasta Upload file [.txt, .xml, .FASTA]

Sequence Number [ID]	Description
1	>HM118516.1 Uncultured bacterium hypothetical protein gene, partial cds; hypothetical protein and Est1 [est1] genes, complete cds; and putative DNA polymerase I [polA] gene, partial cds

Navigation: << 1 >>

Sequence Number [ID]	Sequence Name
1	

Molecule Type*

Please select the sequence molecule type for the sequence that you would like to import.

Organism name*	Qualifier Molecule Type*

Check to save description as a note.

Cancel Import sequence

图 73: 导入序列, FASTA 格式

3) 要完成操作, 用户应单击蓝色的“Import Sequence”按钮。

导入后, 该工具将导航到“Import Report”视图, 参见[错误!未找到引用源。](#)。

插入序列

要将一个序列插入到序列表的特定位置, 用户必须单击序列节顶部的“Insert Sequence”按钮 (在图 74 中显示)。

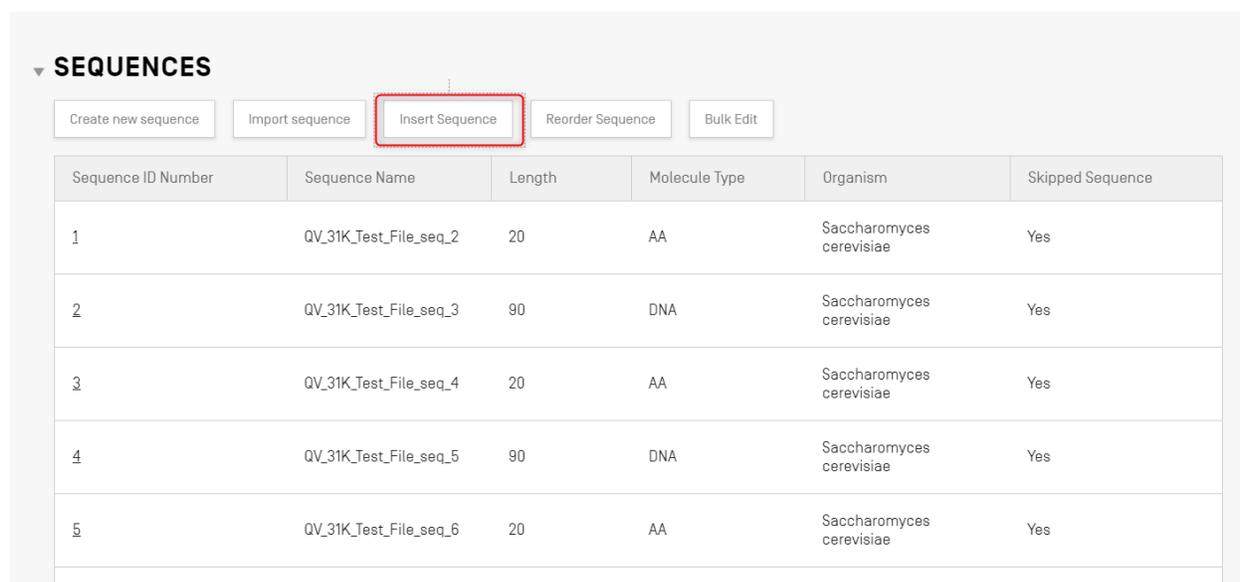


图 74: 插入序列第 1 步

一个带有面板的叠加层将出现（见图 75）。用户必须填写创建序列所需的所有信息（参见**错误!未找到引用源。**）。此外，在面板的左上方，用户必须输入序列在序列表中出现的位置（在图 75 中显示）。

Sequence Number (ID)*
3

Sequence Name
inserted sequence

Molecule Type*
DNA

Residues*
gttagaatgtaggaatgtaggaatctacatgtaggaatgtaggaatgtaggaatctacatgtaggaatgtaggaatgtaggaatctacatgtaggaatgtaggaatgtaggaatctacatgtaggaatgtaggaatgtaggaatctacatgtaggaatgtaggaatgtaggaatctacatgtaggaatgtaggaatgtaggaatctacat

Organism name *
Wohlfahrtiopsis bishoppi

Qualifier Molecule Type
genomic DNA

Mark as an intentionally skipped sequence
 The sequence contains both DNA & RNA fragments

Cancel Insert sequence Insert & Display Sequence

图 75: 插入序列编辑面板

要完成操作，用户可以单击插入“Insert sequence”或“Insert & Display Sequence”，也在图 75 中显示。

SEQUENCES

Create new sequence Import sequence Insert Sequence Reorder Sequence Bulk Edit

Sequence ID Number	Sequence Name	Length	Molecule Type	Organism	Skipped Sequence
1	QV_31K_Test_File_seq_2	20	AA	Saccharomyces cerevisiae	Yes
2	QV_31K_Test_File_seq_3	90	DNA	Saccharomyces cerevisiae	Yes
3	inserted sequence	210	DNA	Wohlfahrtiopsis bishoppi	No
4	QV_31K_Test_File_seq_4	20	AA	Saccharomyces cerevisiae	Yes
5	QV_31K_Test_File_seq_5	90	DNA	Saccharomyces cerevisiae	Yes

图 76: 序列已插入到特定位置

如果用户单击蓝色的“Insert & Display sequence”按钮，在创建序列后就会打开一个可折叠的序列表格，位于图 76 所示项目详细信息（Project Detail）视图下序列表的下方。

重排序列

用户可以通过以下步骤重新组织序列在项目提供的序列表中以什么顺序出现。图 77（开始）到图 79（结果）显示了序列顺序的转换。

- 1) 单击“Reorder Sequence”按钮，如图 77 所示。

SEQUENCES

Create new sequence Import sequence Insert Sequence Reorder Sequence Bulk Edit

Sequence ID Number	Sequence Name	Length	Molecule Type	Organism	Skipped Sequence
1	QV_31K_Test_File_seq_2	20	AA	Saccharomyces cerevisiae	Yes
2	QV_31K_Test_File_seq_3	90	DNA	Saccharomyces cerevisiae	Yes
3	inserted sequence	256	DNA	Wohlfahrtiopsis bishoppi	No
4	QV_31K_Test_File_seq_4	20	AA	Saccharomyces cerevisiae	Yes
5	QV_31K_Test_File_seq_5	90	DNA	Saccharomyces cerevisiae	Yes
6	QV_31K_Test_File_seq_6	20	AA	Saccharomyces cerevisiae	Yes

图 77: 重排序列

2) 选择要移动的序列，并对照当前序列表顺序选择放置位置，如图 78 所示。

20190509-1458 Artificial sequence
S00001_seq_SEQ ID NO 15

1 2 3 4 5

Sequence Number (ID)*
3

New Sequence Position
2

To select a range use "-" between the ids

Cancel Reorder Sequence

图 78: 重排序列编辑面板

3) 现在，该序列出现在图 79 的第二个位置。

SEQUENCES

Create new sequence Import sequence Insert Sequence Reorder Sequence Bulk Edit

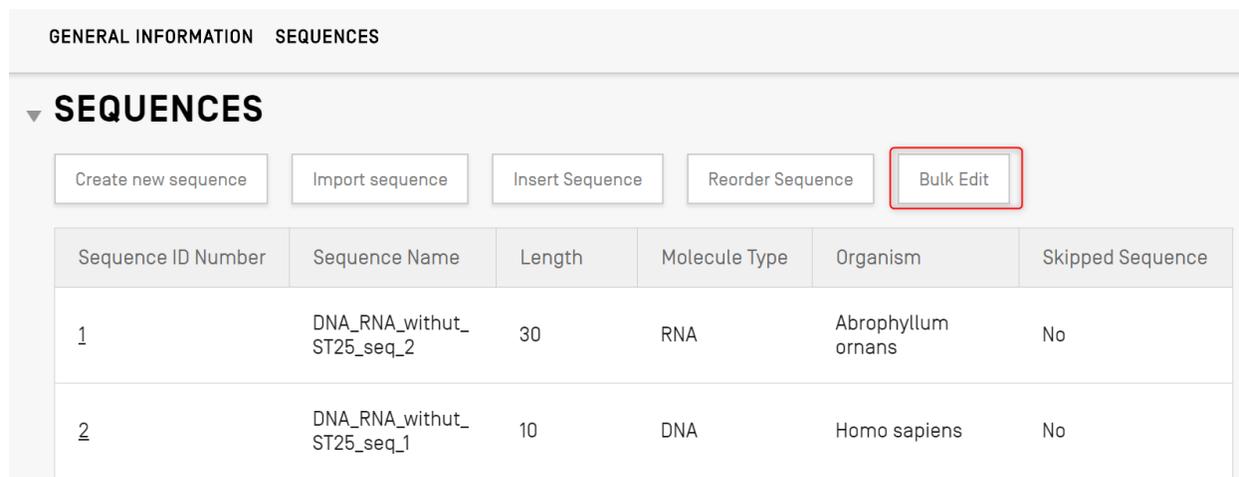
Sequence ID Number	Sequence Name	Length	Molecule Type	Organism	Skipped Sequence
1	QV_31K_Test_File_seq_2	20	AA	Saccharomyces cerevisiae	Yes
2	inserted sequence	256	DNA	Wohlfahrtiopsis bishoppi	No
3	QV_31K_Test_File_seq_3	90	DNA	Saccharomyces cerevisiae	Yes
4	QV_31K_Test_File_seq_4	20	AA	Saccharomyces cerevisiae	Yes
5	QV_31K_Test_File_seq_5	90	DNA	Saccharomyces cerevisiae	Yes

图 79: 重排后的序列

批量编辑

若需要对多个序列进行修改，用户可以使用批量编辑。虽然可以对序列单个进行编辑，但对于有大量序列的项目而言是不可行的。

1) 点击“Bulk edit”，如错误!未找到引用源。所示：



The screenshot shows a web interface for managing sequences. At the top, there are two tabs: 'GENERAL INFORMATION' and 'SEQUENCES'. Below the 'SEQUENCES' tab, there is a section titled 'SEQUENCES' with a dropdown arrow. Underneath, there are five buttons: 'Create new sequence', 'Import sequence', 'Insert Sequence', 'Reorder Sequence', and 'Bulk Edit'. The 'Bulk Edit' button is highlighted with a red rectangular border. Below the buttons is a table with the following data:

Sequence ID Number	Sequence Name	Length	Molecule Type	Organism	Skipped Sequence
1	DNA_RNA_withut_ST25_seq_2	30	RNA	Abrophyllum ornans	No
2	DNA_RNA_withut_ST25_seq_1	10	DNA	Homo sapiens	No

图 80: 批量编辑第 1 步

- 2) 选择 “Type of bulk edit” 为 ‘Qualifier molecule type’ 、 ‘Organism’ 或 ‘Feature’ :

The screenshot shows a web form for bulk editing. The 'Type of bulk edit*' dropdown is set to 'Qualifier molecule type'. Below it, a section titled 'Select the type of bulk edit to continue' contains instructions: 'Use commas to separate individual sequences (for example: "2,5,8"); use a dash to indicate a range of sequences (for example: "2-8").' The 'Molecule Type*' dropdown is set to 'DNA'. A yellow callout box states: 'Only DNA/RNA sequences allow editing of the qualifier 'mol_type' because for amino acid sequences this qualifier value is automatically set to 'protein''. Below this, a note says: 'The qualifier 'mol_type' of hybrid DNA/RNA sequences, where the 'combinedIndicator' = 'Yes', cannot be bulk edited.' At the bottom, there is a text input field for entering DNA sequences.

图 81: 批量编辑第 2 步

正如图 81 所示，系统通知用户，只有核酸序列可以编辑限定符 ‘mol_type’ 的值（因为系统自动将氨基酸序列的相同值设置为 ‘protein’ ）。

- 3) 选择分子类型后，系统提示用户选择批量编辑所适用的核酸序列类型

The screenshot shows the same bulk edit interface as Figure 81, but with the 'Qualifier Molecule Type*' dropdown set to 'unassigned DNA'. The 'Molecule Type*' dropdown remains set to 'DNA'. The table below the form shows a list of sequences with the 'Qualifier molecule type' column highlighted in red.

Sequence ID Number	Sequence Name	Length	Molecule Type	Organism	Qualifier molecule type
3	500 sequences_seq_3	15	DNA	Vaccaria hispanica	genomic DNA

图 82: 批量编辑，识别出 mol_type 限定符

如图 82 所示，系统还警告用户，生物体=“synthetic sequence”的序列的限定符“mol_type”必须是“other DNA”或“other RNA”，如果改变这些值，项目验证时将产生错误。

4) 系统显示待批量编辑的序列表以供选择

Sequence ID Number	Sequence Name	Length	Molecule Type	Organism	Qualifier molecule type
3	500 sequences_seq_3	15	DNA	Vaccaria hispanica	unassigned DNA
7	500 sequences_seq_7	16	DNA	Wabasso hilairoides	unassigned DNA

图 83: 批量编辑，mol_type 限定符得到更新

若批量编辑的类型是生物体：

1) **错误!未找到引用源。**所示，选择“Type of bulk edit”为 Organism (1)。用户必须输入要编辑的 SEQ ID 号范围 (2)。例如，如果用户选择将生物体的值改为“synthetic construct”，系统将通知用户，“mol_type”限定符将根据分子类型自动改为“other DNA”或“other RNA” (3)。

Type of bulk edit*
Organism

Select Range of Sequence IDs*
1-3

Select the type of bulk edit to continue

Organism name*
synthetic construct

The qualifier 'mol_type' value will be automatically changed to 'other DNA' or 'other RNA'.

Sequence ID Number	Sequence Name	Length	Molecule Type	Organism	Qualifier molecule type
1	Seq_1	192	DNA	Mus musculus	genomic DNA
2	Seq_2	20	DNA	Homo sapiens	other DNA
3	Seq_3	20	DNA	Gabaza connectens	other DNA

图 84: 批量编辑生物体第 1 步

若批量编辑的类型是特征：

1) 选择“Type of bulk edit”为 Feature (见**错误!未找到引用源。**) (1)。用户接着必须说明是要编辑现有特征还是要添加新特征 (2)。用户必须输入“Molecule Type” (3) 和将要编

辑的序列 ID 范围 (4)。在用户选择“Edit feature” (2) 的情况下，如果用户选择了将特征位置的值改为“complement(join(1..30,61..90))”，该工具将把 SEQ ID 5 至 8 的所有现有 CDS 特征 (5) 更新为给定值。

Sequence ID Number	Sequence Name	Length	Molecule Type	Organism	Qualifier molecule type
1	QV_31K_Test_File_seq_2	20	AA	Saccharomyces cerevisiae	protein
2	inserted sequence	256	DNA	Wohlfahrtiopsis bishoppi	genomic DNA
3	QV_31K_Test_File_seq_3	90	DNA	Saccharomyces cerevisiae	genomic DNA
4	QV_31K_Test_File_seq_4	20	AA	Saccharomyces cerevisiae	protein
5	QV_31K_Test_File_seq_5	90	DNA	Saccharomyces cerevisiae	genomic DNA
6	Seq_9	26	DNA	Eacles sp.	other DNA

图 85: 批量编辑特征第 1 步

2) 如果用户选择批量特征编辑的类型为“Add feature” (见错误!未找到引用源。): 一个特征位置为“1...13” (5) 的新特征“CHAIN” (4), 将被添加到每个选定的 SEQ ID 1,2,7 (3)。

Sequence ID Number	Sequence Name	Length	Molecule Type	Organism	Qualifier molecule type
1	40000 sequences_seq_1	15	AA	Gabara gigantea	protein
2	40000 sequences_seq_2	16	AA	Daboia russelii	protein
7	40000 sequences_seq_7	15	AA	Wabasso hilairoides	protein

图 86: 批量编辑特征: 添加特征

3) 如果用户选择的批量编辑的类型为“Bulk Skip”（见**错误!未找到引用源。**），要跳过的序列范围被确定为一组 SEQ ID（**1**）。

图 87: 批量编辑特征: 批量跳过

编辑序列

要编辑序列，用户必须单击要修改的相应序列的序列 ID 号，如图 88 中显示（序列 ID‘1’）。

注：

所有符合产权组织标准 ST. 26 的序列列表必须以序列 ID 号 1 开始，并连续编号至序列长度。

SEQUENCES

Create new sequence Import sequence Insert Sequence Reorder Sequence Bulk Edit

Sequence ID Number	Sequence Name	Length	Molecule Type	Organism	Skipped Sequence
1	DNA_RNA_withut_ST25_seq_2	30	RNA	Abrophyllum ornans	No
<u>2</u>	DNA_RNA_withut_ST25_seq_1	10	DNA	Homo sapiens	No

图 88: 编辑序列

这将在项目详细信息视图中打开一个新节，位于序列节的下方，如图 89 所示。

该序列节由排第一的突出显示的“Basic Information”编辑面板（**1**）、排第二的“Features”列表表格（**2**）和排最后的“Sequence”残基编辑面板（**3**）组成。

The screenshot displays the 'SEQUENCE 1' editing interface. It is divided into three main sections:

- SEQUENCE 1:** A summary box containing:
 - Sequence Number (ID): 1
 - Sequence Name: ST25_221_BINDING_edited_seq_1
 - Length: 10
 - Molecule Type: AA
 - Organism: synthetic construct
- FEATURES:** A table for defining features. It includes an 'Add feature' button and a table with columns for Feature Key, Location, and Qualifiers.

Feature Key	Location	Qualifiers
source	1..10	mol_type = protein organism = synthetic construct
- SEQUENCE:** A text area for the sequence itself, showing 'ACCEWREXXXX' and a length of 10.

图 89: 编辑序列、编辑面板

特征

根据 WIPO ST.26, 每个序列必须具备至少一个与之相关的特征: ‘source’, 取决于分子类型。每个 source 特征都必须有两个强制限定符: ‘organism’ 和 ‘mol_type’。

特征表格有三列: 特征键、特征在遗传序列中的位置以及与单个序列特征相关的限定符。

特征位置指明特征存在于序列的哪个片段中。WIPO ST.26 规定了指明特征位置时所允许的格式, 具体如下:

- 单一残基编号: x
- 界定序列范围的残基号: x..y
- 位于第一个指定残基号之前或位于最后一个指定残基号之后的残基: <x, >x, <x..y, x..>y, <x..>y
- 两个相邻核苷酸之间的位点: x^y
- 通过链内交叉链接而连接的残基号: x..y

位置运算符可以用来形成复杂的位置说明:

- “**join (location, location, ... location)**”: 多个位置被连接起来 (以端到端的方式放置), 形成一个连续的序列。

- “**order (location, location, ... location)**”：以指明的顺序找到各种元素，但对连接这些元素是否合理不做任何暗示。
- “**complement (location)**”：表示当按 5'至 3'的方向或按模仿 5'至 3'的方向读取时，该特征位于与位置描述符指明的序列段互补的链上。

要在序列中添加新特征，请单击所选序列的特征节中的“Add feature”按钮，如下图 90 所示。

▼ SEQUENCE 2

Sequence Number (ID) 2 Molecule Type DNA
 Sequence Name Inserted sequence Organism Wohlfahrtiopsis bishoppi
 Length 41

▼ FEATURES

Add feature

Feature Key	Location	Qualifiers
source	1..41	mol_type = genomic DNA organism = Wohlfahrtiopsis bishoppi

Navigation: [Home] [Left] [1] [Right] [End]

▼ SEQUENCE

gatagtatgt atatatagta gtatgatgat gatgatgga t 41

图 90: 为序列添加特征

接下来，在打开的叠加层中（如图 91 所示），从特征键下拉列表⁸中选择一个条目，并指明该特征在所适用的序列中的特征位置。

⁸ 参见产权组织标准 ST.26—附件一第 5 节和第 7 节：核苷酸和氨基酸特征键列表。

▼ FEATURES

Add feature

Feature Key	Location	Qualifiers	
source	1..23	mol_type = genomic DNA organism = Tinamus osgoodi	🗑️

1

Feature 1

Feature Key*
D_segment

Feature Location*
4..15

For a detailed explanation of the format of the feature location please [click here](#)

▼ QUALIFIERS

Add qualifier

Cancel Create Feature

图 91：添加特征，叠加层

在这个阶段，也可以将限定符添加到特征中，这将在下一节介绍。

Feature 1

Feature Key*
CDS

Feature Location*
1..13

Automatic addition of a translation qualifier is currently on. This means a separate sequence will be generated and a protein_id qualifier created if the translation qualifier value meets the minimum length requirement.

For a detailed explanation of the format of the feature location please [click here](#)

图 92：对特征位置进行详细解释的弹出链接

增加了一个链接，以对特征位置的格式进行详细解释。该链接位于“click here”这句话的末尾。在**错误!未找到引用源。**中，该链接下方已划线，使用户更容易识别。

要完成操作，单击蓝色的“Create Feature”按钮，如图 91 所示。

CDS 特征

CDS 特征型用于描述蛋白质编码序列⁹。CDS 特征可以选择性地包含所属序列片段的氨基酸翻译，如果满足了最低长度要求，将作为一个独立的序列出现在项目中。在原始序列的 CDS 特征中，已翻译氨基酸序列的 ID 参考号将在 “protein_id” 限定符中提供。

在为序列创建 “CDS” 特征时，‘translation’ 限定符（默认 “Genetic Code” 值为 1–“Standard Code”）可自动添加到 CDS 特征中，限定符的值为特征位置所示的已翻译的序列残基块。相关的 ‘protein_id’ 和单独的氨基酸序列也可以通过勾选项目详细信息页面顶部提供的基本信息（Basic Information）中的勾选框来生成（见图 96）。然而，这个限定符不是必填项，可以在生成后删除。用户也可以手动创建 ‘translation’ 和 ‘protein_id’ 限定符，限定符引用已由用户创建的相关翻译序列 ID。

注：

从 2.1.0 版本开始，“Automatically add a translation qualifier..”复选框默认为勾选。

自动创建 CDS 特征

自动创建 CDS 特征限定符的步骤如下：

- 1) 在特定序列显示中，单击 “Add feature” 按钮，选择 “CDS” 作为特征键。如果勾选了基本信息中的 “automatically add a translation qualifier”，当 CDS 特征被添加到核苷酸序列时，它将自动添加 ‘translation’ 限定符、该限定符的值和 ‘protein_id’ 限定符及其相关独立氨基酸序列（如果合适）。
- 2) 用户也可以选择手动创建 ‘translation’ 限定符。

⁹ <https://www.ddbj.nig.ac.jp/ddbj/cds-e.html>

图 93: 创建翻译限定符

- 3) 当用户完成对特征及其相关限定符的编辑后，必须单击“Create Feature”按钮进行保存。由此产生的 CDS 特征如图 94 所示。

Feature Key	Location	Qualifiers
CDS	1..13	protein_id = 3 translation = MYIN
source	1..237	mol_type organism = test bla

图 94: 特征表中新 CDS 特征

然后，该工具将为项目**创建一个新序列***，该序列具有以下属性：

- **序列 ID 号**=序列 ID 号的下一个可用值
- **长度**=所翻译序列的长度
- **序列名称**=“translation”限定符的“Sequence Name”字段中所给的值（见**错误!未找到引用源。**）。如果没有提供名称，将提供默认的序列名称（‘Seq_#’）。
- **分子类型**=“AA”
- **生物体名称**=与原始序列提供的值相同
- **限定符分子类型**=“protein”
- **序列残基**=原始序列的翻译值

上述过程完成后所产生的结果示例如图 95 所示。

▼ **SEQUENCE 3**

Sequence Number (ID) 3 Sequence Name Seq_3 Length 4	Molecule Type AA Organism Caballerocotyla klawei	
---	---	--

▼ **FEATURES**

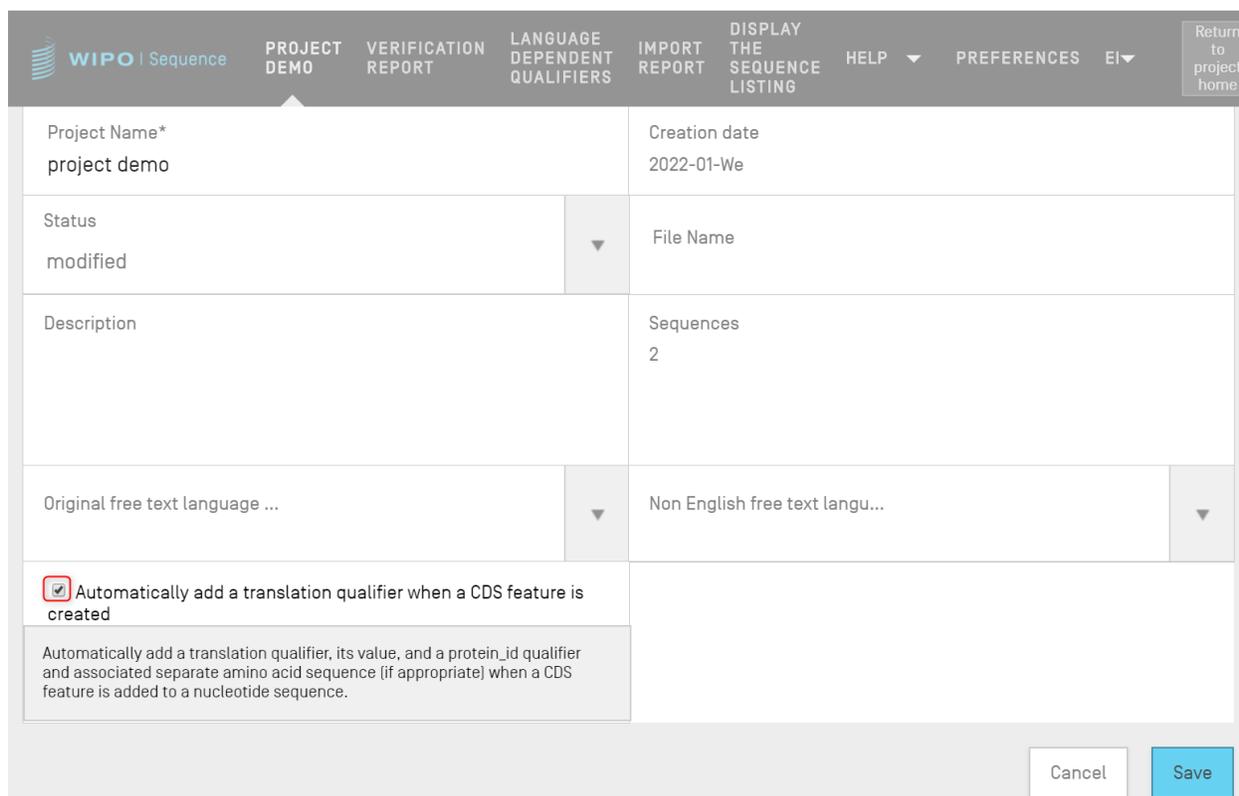
Feature Key	Location	Qualifiers	
COILED	1..4	note = qual val	
source	1..4	mol_type = protein organism = Caballerocotyla klawei	

图 95: CDS 特征, 生成的氨基酸序列

注, 关于翻译序列的创建:

只有当独立的翻译序列至少有 4 个具体定义的残基 (例如 AXTG 算 3 个字符) 时才能被创建。

在修改 “translation” 限定符时, 如果限定符的值减少到 4 个具体定义的残基以下, 那么相关的序列译文将被删除, ‘protein_id’ 限定符也会被删除。



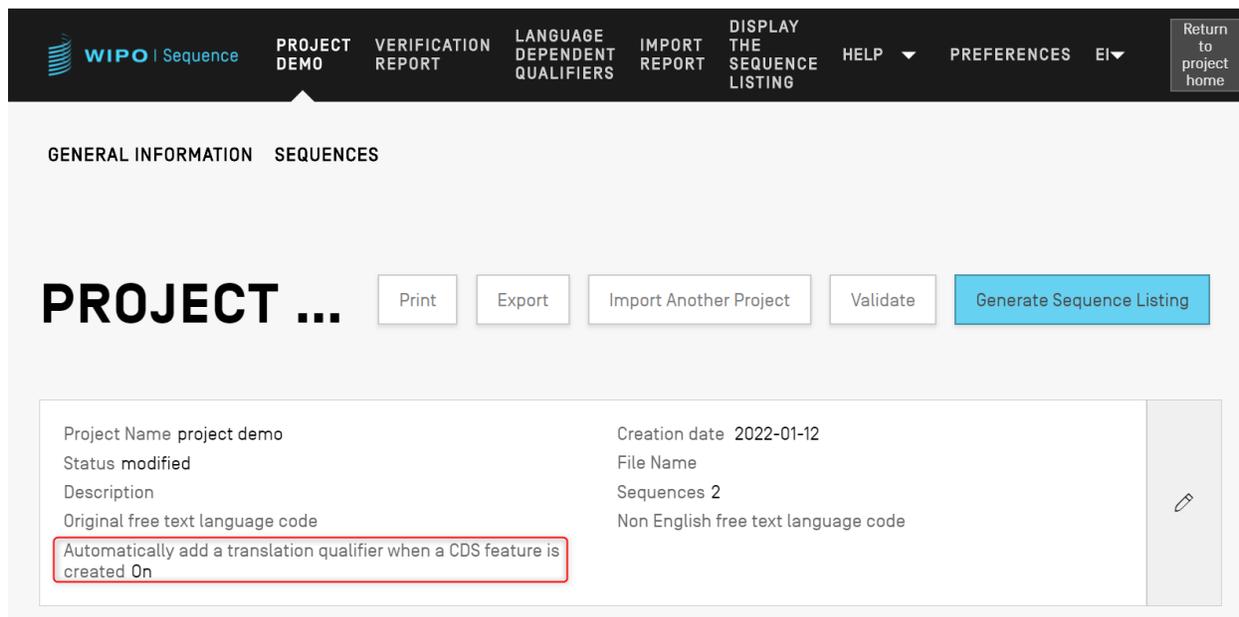
WIPO Sequence		PROJECT DEMO	VERIFICATION REPORT	LANGUAGE DEPENDENT QUALIFIERS	IMPORT REPORT	DISPLAY THE SEQUENCE LISTING	HELP	PREFERENCES	EI	Return to project home
Project Name*		Creation date								
project demo		2022-01-We								
Status			File Name							
modified										
Description		Sequences								
		2								
Original free text language ...			Non English free text langu...							
<input checked="" type="checkbox"/> Automatically add a translation qualifier when a CDS feature is created										
Automatically add a translation qualifier, its value, and a protein_id qualifier and associated separate amino acid sequence (if appropriate) when a CDS feature is added to a nucleotide sequence.										
<input type="button" value="Cancel"/> <input type="button" value="Save"/>										

图 96: : 翻译限定符自动生成复选框：取消选择以关闭

CDS 特征包含 pseudo 或 pseudogene 限定符时的建议：

确保在向 CDS 特征添加 pseudo 或 pseudogene 限定符时关闭自动翻译。

如果在向 CDS 特征添加 pseudo 或 pseudogene 限定符时未关闭自动翻译（如果勾选了**错误!未找到引用源。**所示的复选框），则在更新 CDS 特征时，将自动添加翻译限定符。要更正该错误，请关闭项目的自动翻译，然后打开 CDS 特征，删除翻译和 protein_id 限定符，再更新特征。



The screenshot shows the 'PROJECT ...' page in the WIPO Sequence application. The top navigation bar includes 'WIPO | Sequence', 'PROJECT DEMO', 'VERIFICATION REPORT', 'LANGUAGE DEPENDENT QUALIFIERS', 'IMPORT REPORT', 'DISPLAY THE SEQUENCE LISTING', 'HELP', 'PREFERENCES', and 'Return to project home'. Below the navigation bar, there are tabs for 'GENERAL INFORMATION' and 'SEQUENCES'. The main content area displays the project name 'project demo' and various fields: 'Creation date 2022-01-12', 'File Name', 'Sequences 2', 'Original free text language code', and 'Non English free text language code'. A checkbox labeled 'Automatically add a translation qualifier when a CDS feature is created' is checked and highlighted with a red box. There are also buttons for 'Print', 'Export', 'Import Another Project', 'Validate', and 'Generate Sequence Listing'.

图 97: 翻译限定符自动生成: 已开启

如果用户希望自动生成翻译限定符，可以从限定符编辑面板中设置翻译表值和序列名称。用户创建特征时，该工具将执行翻译，然后为特征添加“protein_id”限定符和带有翻译值的新序列。

只有当特征位置或限定符“transl_table”“transl_except”或“codon_start”之一改变其值时，才会再次执行翻译，在这种情况下，链接的序列将被更新。

注：

如果翻译值改变，链接的序列将自动更新其值。但是，如果链接的核苷酸序列被修改，则翻译限定符的值不会改变。如果“protein_id”限定符在创建后被修改，则链接的序列将失去与原始序列的关联。

关于使用终止密码子的建议：

通常，终止密码子只应出现在 CDS 特征的末端，表示编码氨基酸序列的结束。除非有‘transl_except’限定符，表明该终止密码子将被翻译为特定氨基酸，否则绝不能出现在 CDS 特征的中间。

如果在 CDS 特征的中间发现了终止密码子（下方以黄色突出显示），且没有‘transl_except’限定符表明该终止密码子将被翻译为特定氨基酸，如图 98 所示，那么该工具应该在该点停止翻译，并显示红色横幅，告知用户将不会生成翻译。

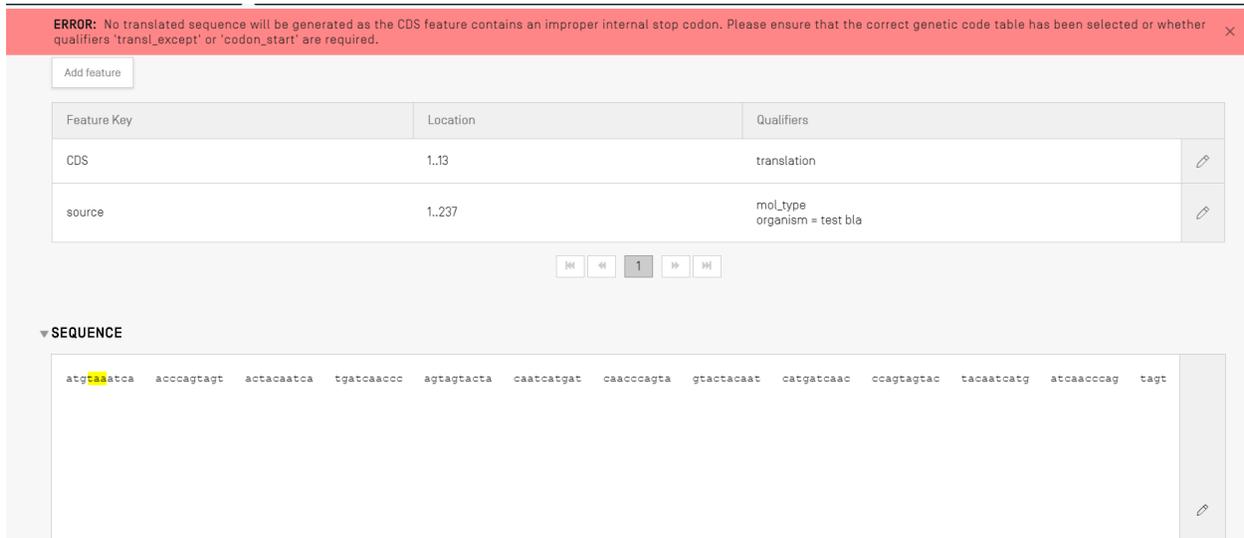


图 98: 翻译限定符自动生成复选框已开启，序列中包含错误的终止密码子

验证报告中会列出错误，以提示用户其编码序列中存在问题，如图 99 所示。

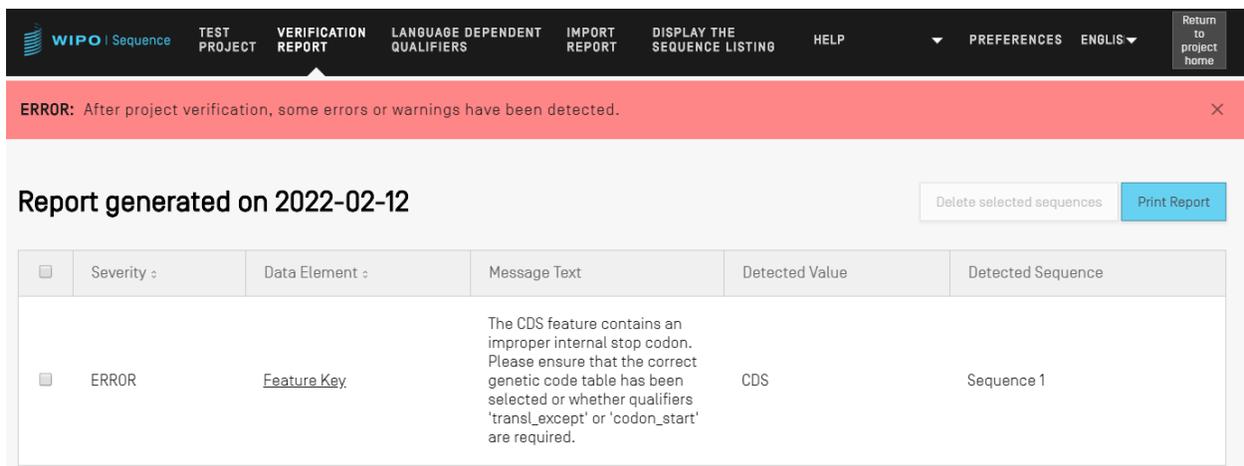


图 99: 错误的内部终止密码子 CDS 特征错误

限定符

要查看一个特征的限定符，用户必须首先从相关序列的特征表中选择一个特征。在图 100 的示例中，“source”特征的铅笔图标被突出显示。

GENERAL INFORMATION SEQUENCES

3 Seq_3 4 AA Caballerocotyla klawei No

1

▼ SEQUENCE 2

Sequence Number (ID) 2 Molecule Type DNA
 Sequence Name ST25_221_BINDING_edited_seq_2 Organism Caballerocotyla biparastica
 Length 237

▼ FEATURES

Add feature

Feature Key	Location	Qualifiers
CDS	1..13	protein_id = 3 translation = MYIN
source	1..237	mol_type organism = Caballerocotyla biparastica

1

图 100: 添加/编辑限定符第 1 步

单击图 101 中突出显示的铅笔按钮，将打开以下的特征叠加层，如**错误!未找到引用源。**所示。

Feature 2

Feature Key	Feature Location*
source	1..237

For a detailed explanation of the format of the feature location please [click here](#)

▼ QUALIFIERS

Add qualifier (2)

Qualifier Name	Qualifier Value
mol_type	
organism	Caballerocotyla biparastica

Cancel Update feature

图 101: 添加/编辑限定符第 2 步

通过单击每行右侧的铅笔图标，可以编辑现有的限定符（1），用户也可以通过单击“Add qualifier”按钮（2）为当前选中的特征添加新的限定符。

在编辑或添加限定符时，用户将看到图 102 所示的两个字段：‘Qualifier name’¹⁰（从下拉列表中选择）和‘Qualifier value’。

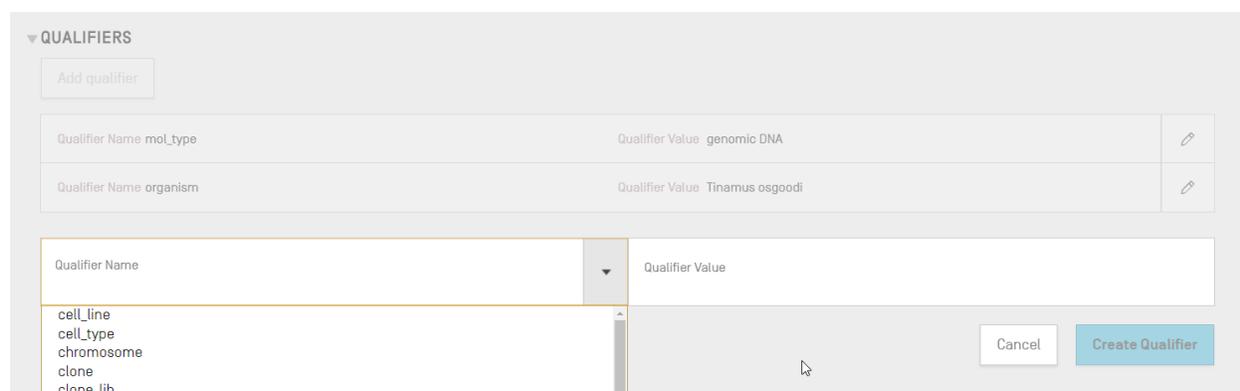


图 102: 限定符编辑面板

根据限定符的类型，限定符值字段会有不同的行为：

- **具有预定义值的限定符。**值字段是一个下拉字段，用户可以为限定符选择一个预定义值，如下**错误!未找到引用源。**中的样例所示：

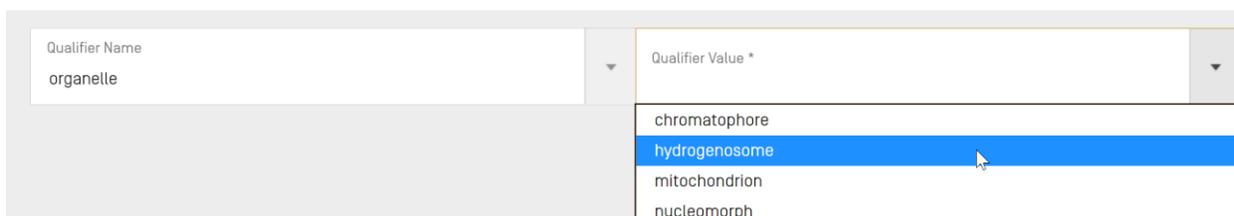


图 103 具有预定义值的限定符

- **带有自由文本的限定符。**值字段是自由文本字段。除限定符名称和限定符值（值仅可为英语）之外，会出现两个附加字段，允许用户提供两种语言代码（例如“ru”）以及在非英语限定符值中提供相应语言的值，如**错误!未找到引用源。**中的下例所示：

¹⁰ 参见产权组织标准 ST.26-附件一第 6 节和第 8 节：核苷酸和氨基酸所有可能的限定符。

Qualifier Name *	Qualifier Value
note	NOTE: this value may require translation for National/Regional procedures.
Language*	Non English Qualifier Value
ru - Russian	

Cancel Create Qualifier

图 104: 带有自由文本值的限定符

语言代码字段与项目详细信息中的“Non English Free text language code”（非英语自由文本语言代码）分配到相同的值。

用户可以通过手动输入或从 XLIFF 文件中导入适当相关语言，为每种选定的语言提供一系列非英语值。

- **具有预定义格式的限定符。**值字段是自由文本字段，但输入的值须经过验证，以确保其符合产权组织 ST.26 附件一第 6 节¹¹所规定的具体规则。在**错误!未找到引用源。**的下例中，日期没有以正确的格式提供：

Qualifier Name	Qualifier Value
collection_date	2002-08-

Invalid qualifier value format. The value for the qualifier 'collection_date' must have the format 'YYYY-MM-DD', 'YYYY-MM' or 'YYYY'.

图 105: 具有预定义格式的限定符

- **不允许带值的限定符。**限定符值字段不可编辑，如**错误!未找到引用源。**中的下例所示：

Qualifier Name	Qualifier Value
germline	Qualifier Value

图 106: 不允许带值的限定符

注：

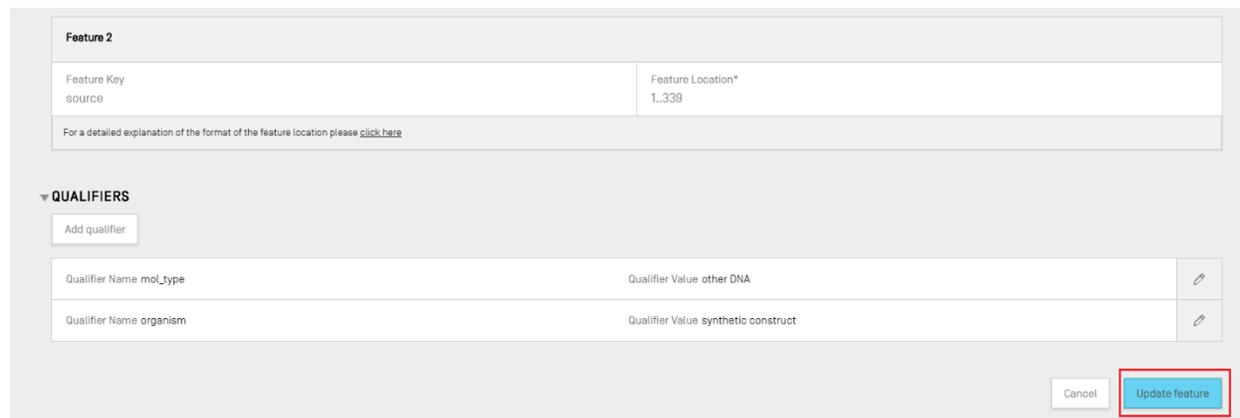
注：

¹¹ 参见产权组织标准 ST.26 – 附件一第 6 节：关于这些限定符的规则。

创建后，`source` 特征的特征位置无法编辑，`mol_type` 和 `organism` 限定符只能编辑不能删除。

完成后，用户必须单击蓝色的“**Create Qualifier**”按钮，以添加新创建的限定符，或者单击“**Save**”以保存对现有限定符作出的修改。

最后一步，一旦添加/修改了限定符，用户必须单击特征叠加层底部的“**Update feature**”按钮以继续，如图 107 所示。



The screenshot shows a web interface for editing a feature. At the top, there is a section titled "Feature 2" with a table containing "Feature Key" (source) and "Feature Location*" (1..339). Below this is a link for a detailed explanation of the feature location format. The main section is titled "QUALIFIERS" and contains an "Add qualifier" button. Below the button is a table with two rows of qualifiers:

Qualifier Name	Qualifier Value	Action
<code>mol_type</code>	other DNA	✎
<code>organism</code>	synthetic construct	✎

At the bottom right of the interface, there are two buttons: "Cancel" and "Update feature". The "Update feature" button is highlighted with a red box.

图 107: 编辑限定符后更新特征

验证报告

要打开项目序列表的验证报告，在项目详细信息视图中，用户可以单击视图顶部菜单栏中的“**VERIFICATION REPORT**”链接。

关于如何生成验证报告的进一步详情，参见**错误!未找到引用源。**。

WIPO | Sequence PROJECT DEMO **VERIFICATION REPORT** LANGUAGE DEPENDENT QUALIFIERS IMPORT REPORT DISPLAY THE SEQUENCE LISTING HELP ▾ PREFERENCES EI ▾ Return to project home

GENERAL INFORMATION SEQUENCES

PROJECT ...

Print Export Import Another Project Validate Generate Sequence Listing

Project Name project demo	Creation date 2022-01-12	✎
Status modified	File Name	
Description	Sequences 2	
Original free text language code	Non English free text language code	
Automatically add a translation qualifier when a CDS feature is created On		

图 108: 查看验证报告第 1 步

根据项目序列列表是否有效，将产生两个视图：图，产生了错误/警告；图 110，经过了成功验证。

WIPO | Sequence PROJECT DEMO **VERIFICATION REPORT** LANGUAGE DEPENDENT QUALIFIERS IMPORT REPORT DISPLAY THE SEQUENCE LISTING HELP ▾ PREFERENCES EI ▾ Return to project home

Report generated on 2022-01-12

Delete selected sequences Print Report

<input type="checkbox"/>	Severity ▾	Data Element ▾	Message Text	Detected Value	Detected Sequence
	ERROR	Application Identification	The mandatory IP Office Code is missing.	-	
	WARNING	Earliest Priority Application Identifications	Priority application information has been entered, but no prior application has been designated as the earliest. The Earliest priority application must be designated when a priority claim is made to an earlier application.	-	

图 109: 验证报告样例，错误/警告

单击图 109 右上角所示的“Print Report”按钮，可以将验证报告导出为 PDF 格式。生成的报告将显示在图 9 所示的 PDF 查看器中。



图 110 验证报告，无错误/警告

语种相关限定符

本节详细介绍语种相关限定符视图中提供的各种功能。在项目中允许使用“free text”值的限定符在项目页面的“LANGUAGE DEPENDENT QUALIFIERS”视图中进一步引用。

每当在当前项目中添加一个语种相关限定符时，该限定符也会在上述视图中显示，如图 111 所示。

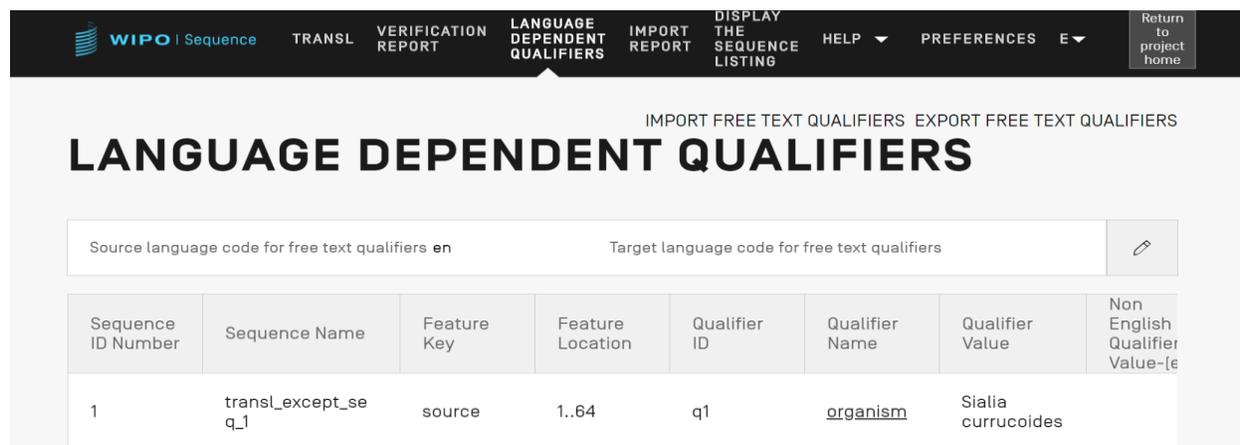


图 111：语种相关限定符视图

用户可以通过单击“Qualifier Name”来修改限定符相关的翻译后的自由文本值，如图 111 所示，这将在该表下面打开一个带有编辑面板的叠加层。

用户需要为自由文本限定符 XLIFF 文件的导出提供源语言代码和目标语言代码，其中翻译值需要在重新导入 XLIFF 文件前由译员提供。

导入自由文本限定符

如果用户点击“IMPORT FREE TEXT QUALIFIERS”按钮，该工具将打开文件资源管理器，使用户能够进行浏览，找到并选择想要导入的（.XLIFF）文件。提供了多个验证步骤以确保在源语言和目标语言值之间进行正确的映射。

所选文件必须包含以下数据项目：

- 项目名称
- 目标语言代码
- 源语言代码
- 各 XLIFF 单位元素：
 - 限定符的唯一 ID（参考格式：字母 q 后加数字）
 - 源语言标签中的限定符值
 - 目标语言标签中的限定符值

```
<xliff xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:schemaLocation="urn:oasis:names:tc:xliff:document:1.2 http://docs.oasis-open.org/xliff/v1.2/os/xliff-core">
  <file original="basic1" datatype="plaintext" source-language="en" target-language="fr">
    <body>
      <trans-unit id="q1" sequenceIDNumber="1" sequenceName="basic_seq_1" featureKey="source" featureLocation="1..26" qualifierName="organism">
        <source>Una virus</source>
        <target>organism fr</target>
      </trans-unit>
      <trans-unit id="q2" sequenceIDNumber="2" sequenceName="basic_seq_2" featureKey="source" featureLocation="1..26" qualifierName="organism">
        <source>Una virus</source>
        <target>organism fr</target>
      </trans-unit>
    </body>
  </file>
</xliff>
```

图 112：有效 XLIFF 文件样例

一旦用户确认选择了要导入的文件，该工具将要求用户通过确认一系列验证步骤来验证是否真的要继续下去（见图 113 至图 117）。

- 系统会比较导入文件的项目名称和所选项目的名称：

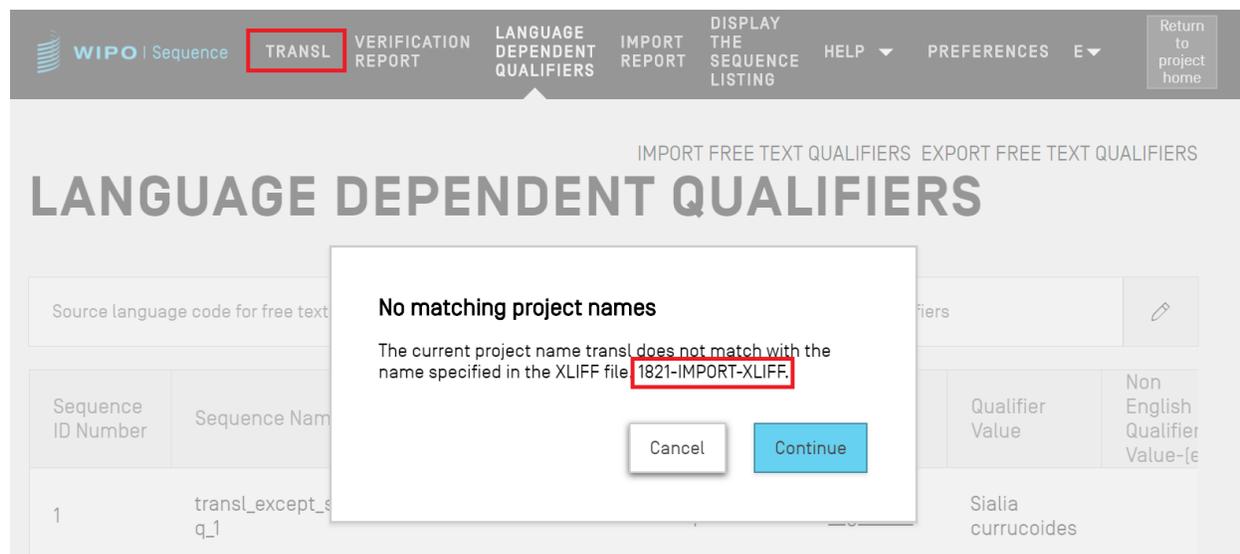


图 113: 项目名称验证

- 如果有任何限定符无法建立映射，系统会通知用户：

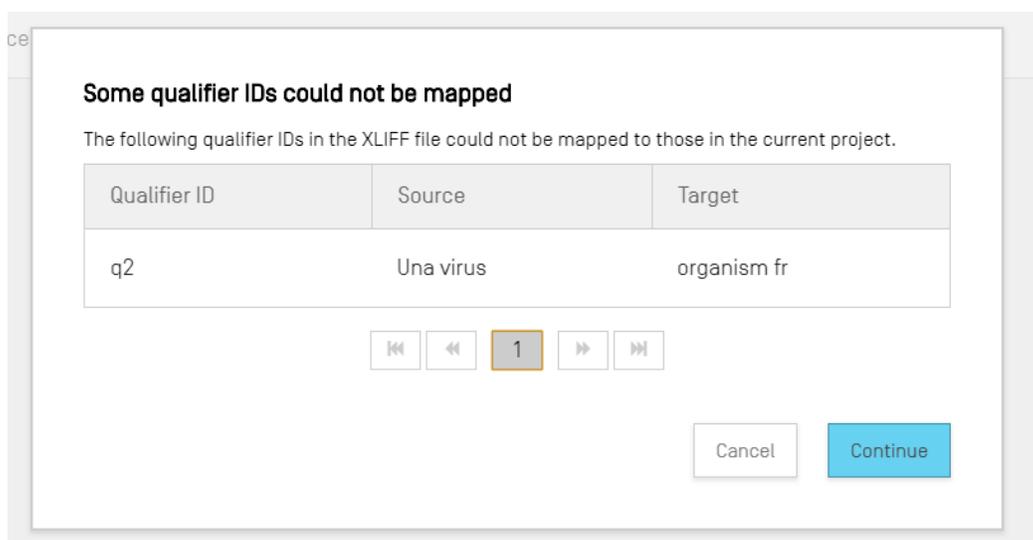


图 114 限定符映射验证步骤

- 系统将通知用户关于源语言和限定符值的变化：

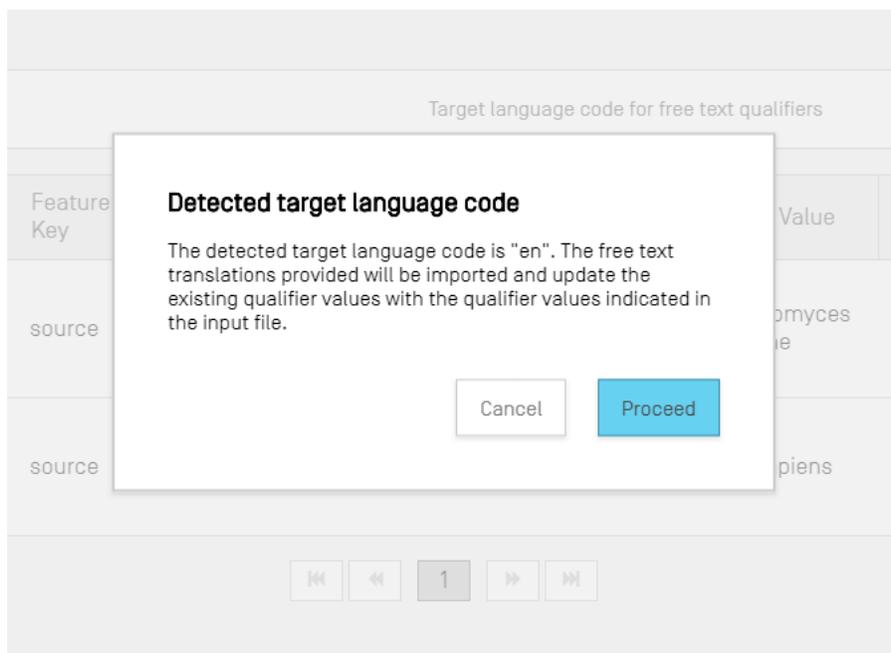


图 115: 源语言验证步骤

- 系统将通知用户关于目标语言和限定符翻译值的变化:

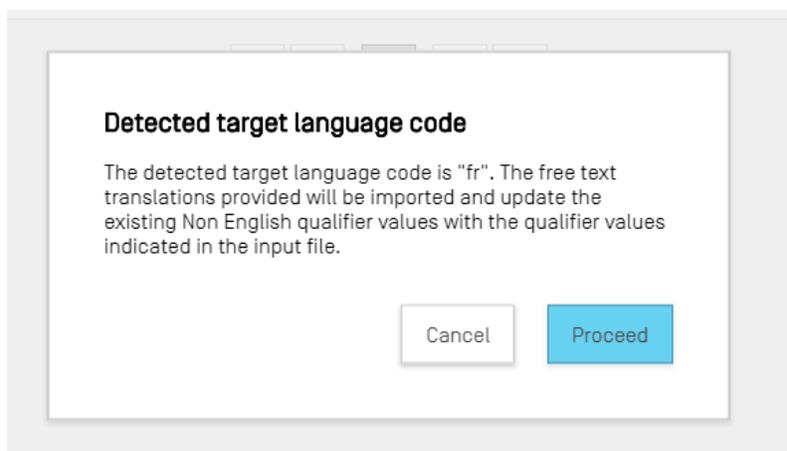


图 116: 目标语言验证步骤

在此之后，会收到上方蓝色横幅“**SUCCESS: THE FREE TEXT QUALIFIER HAS BEEN IMPORTED SUCCESSFULLY**”，以及一份导入报告，其中具体显示语种相关自由文本限定符之前的值和现在导入的值。

IMPORT FREE TEXT QUALIFIERS REPORT

Sequence ID Number	Feature Key	Feature Location	Qualifier ID	Qualifier Name	Original qualifier value	Original Non English qualifier value	Imported qualifier value
2	source	1.19	q3	organism	Sialia currucoides		Construction synthétique
1	source	1.64	q1	organism	Sialia currucoides		

图 117: 导入自由文本限定符报告

用户可以点击“RETURN TO FREE TEXT QUALIFIERS”回到自由文本限定符视图，图 117 圈出了该重点。

导出自由文本限定符

如果用户点击视图顶部的“EXPORT FREE TEXT QUALIFIERS”按钮，然后在对话框中选择保存限定符文本文件的文件名和位置，项目的**所有**自由文本限定符都将被导出并保存为 XLIFF 文件格式。

文件将包括：

- 项目的源语言
- 项目的目标语言
- 自由文本限定符值

- 经翻译的限定符自由文本值¹²
- 图 117 中的表格提供的相关限定符和特征信息

在按照**错误!未找到引用源。** - 125 所示步骤提供适当翻译后，可以在工具中再次查看、编辑和导入该文件。

导入报告

本节详细介绍导入报告视图中提供的各种功能。

如果一个项目是从序列表（ST.25 或 ST.26）中导入的，或者当用户从文件（格式为 ST.26、ST.25、raw、FASTA 或 multi-sequence）中导入多个序列时，那么相应的导入报告将包括一个表，表中包含对导入数据所做的所有变更，以使导入的序列表符合正确的格式，从而纳入项目中。

如果一个项目不是通过导入过程创建的，也没有将任何序列导入到项目中，导入报告视图将显示图 118 所示的横幅。

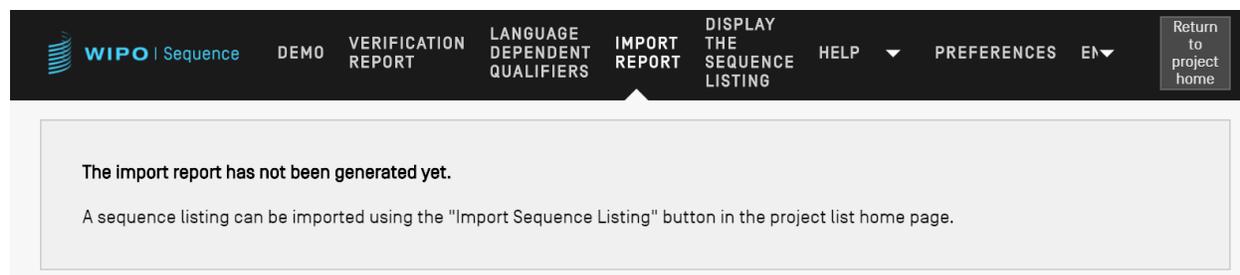


图 118: 导入报告视图，未执行导入

如果导入已成功完成，且项目已创建，则视图将显示以下两个表格：

- 导入报告表（参见**错误!未找到引用源。**）
- 变更数据表（参见**错误!未找到引用源。**）

¹² 导出时项目提供的所选非英语自由文本语言代码的翻译。

导入报告表

Type of Note	Data Element	Message Text	Detected Sequence
Global	<130>	During import, the 'Applicant File Reference' has been automatically assigned the following value: USPTO; this value can be edited within the project.	
Global	<110>	One organization or person must be selected as the first applicant or one person must be selected as the inventor, together with the appropriate language code.	

图 119: 导入报告样例

只有在导入文件产生错误时，才会显示导入报告表，并显示以下几列内容：

- **注释类型：**“INDIVIDUAL”表示与特定序列有关的消息，或者“GLOBAL”泛指一个或多个序列；
- **数据元素代码：**来自源文件，用于 ST.25 序列表；
- **消息文本：**具体的消息，其中包括已查明的有关问题的信息和为纠正该问题所做的更改（如有）；
- **检测到的序列：**与消息有关的导入序列的序列号（当类型为“INDIVIDUAL”时；否则该字段留空）。

变更数据表

Global	<400>	Three-letter amino acid symbols have been replaced with their corresponding one-letter codes.
--------	-------	---

Changed Data

Origin Tag	Origin Element Name	Origin Element Value	Target Element Name	Target Element Value	Transformation	Origin Sequence ID	Sequence ID Number
<221>	Name/Key		Feature Key	misc_feature	The custom feature key has been replaced with a recommended key (see Annex VII, ST.26).	1	1
<223>	Other information		Qualifier Name	note	A 'note' Qualifier has been created.	1	1

图 120: 变更数据表样例

本表显示了在导入过程中经历过转换或变更的任何数据。以下数据以表列形式显示（见图 120）：

- **Origin Tag（原始标签）**：元素类型的数据元素代码，用于 ST.25 序列表；
- **Origin Element Name（原始元素名称）**：元素类型的对应名称；
- **Origin Element Value（原始元素值）**：源文件中原始元素的对应值；
- **Target Element Name（目标元素名称）**：相当于 ST.26 元素名称，是信息将存储于项目中的位置；
- **Target Element Value（目标元素值）**：为项目的目标元素名称设置的值；
- **Transformation（转换）**：对元素所作变更/转换的说明；
- **Sequence ID Number（序列 ID 号）**：项目中被转换元素的相关序列的 ID 号。

显示序列表

本节详细介绍显示序列表视图中提供的各种功能。

WIPO Sequence 允许用户以比 XML 更容易被人阅读的格式生成序列列表。打开“DISPLAY THE SEQUENCE LISTING”视图后，首先会出现一个蓝色的横幅，表示序列列表文件已经成功生成，同时会出现以.html 或.txt 格式显示上述序列列表这两个选项（见图 121）。

有关如何生成序列列表的进一步详情，参见[错误!未找到引用源。](#)。



图 121: 显示序列列表，成功生成

如果没有成功生成给定项目的序列列表，那么显示序列列表视图将禁用“Display Sequence Listing”和“Export Sequence Listing as .txt file”按钮，用户将看到图 122 中所示的错误。

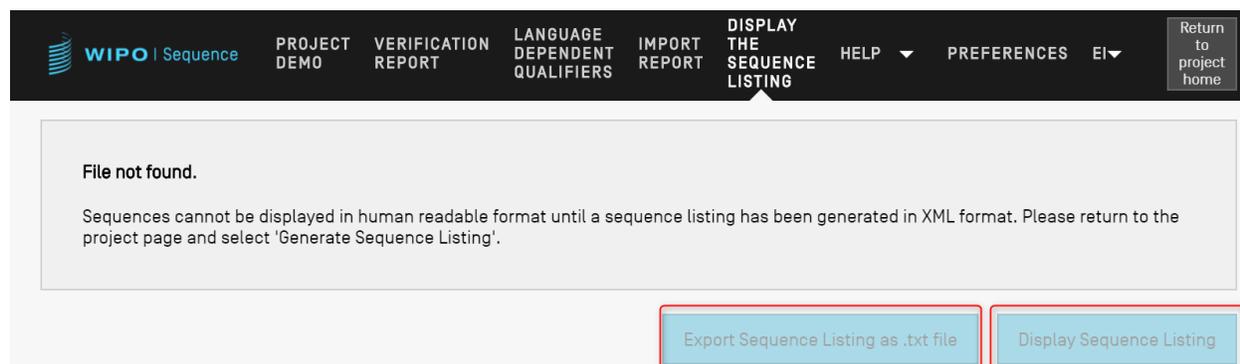
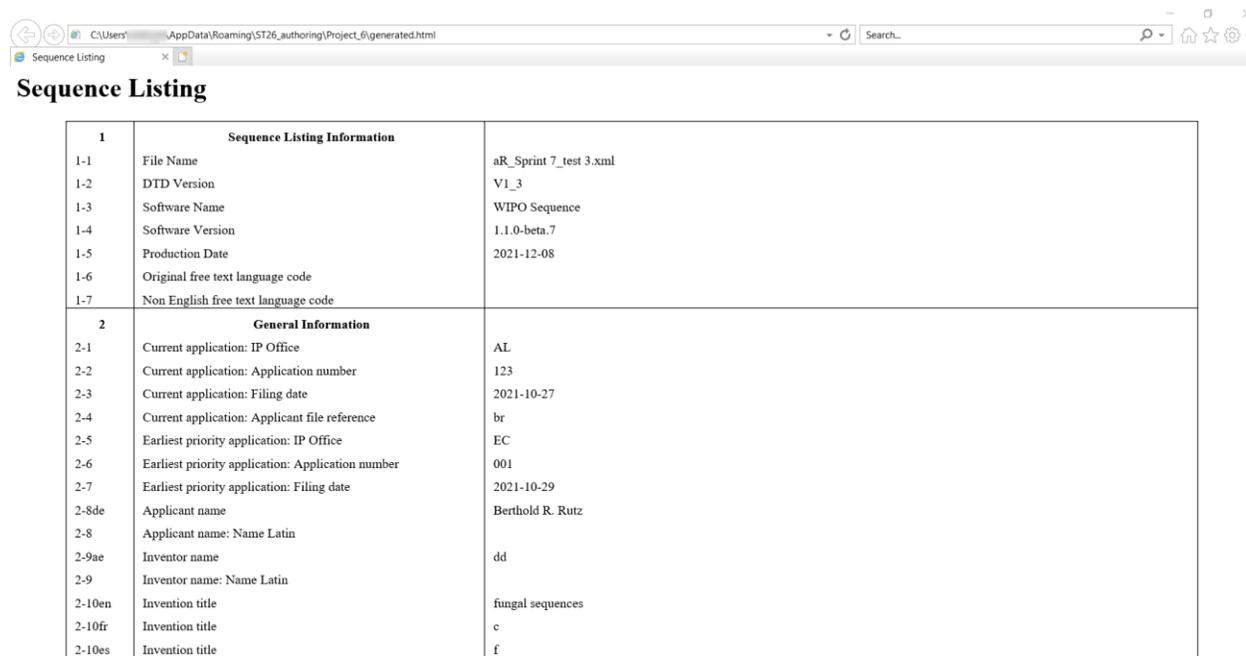


图 122: 显示序列列表，文件未生成

用户单击“Display Sequence Listing”按钮时，一个 HTML 文件将在默认浏览器中打开。这提供了一个符合 ST.26 标准的 XML 文件的格式化视图，以便用户更方便地看到特定字段的值，如图 123 举例所示。

注：

要用另一种语言显示序列列表，必须再次生成序列列表。首先要在常规信息节指出新的非英语自由文本语言代码，然后重复上述步骤。



1		
Sequence Listing Information		
1-1	File Name	aR_Sprint 7_test 3.xml
1-2	DTD Version	V1_3
1-3	Software Name	WIPO Sequence
1-4	Software Version	1.1.0-beta.7
1-5	Production Date	2021-12-08
1-6	Original free text language code	
1-7	Non English free text language code	
2		
General Information		
2-1	Current application: IP Office	AL
2-2	Current application: Application number	123
2-3	Current application: Filing date	2021-10-27
2-4	Current application: Applicant file reference	br
2-5	Earliest priority application: IP Office	EC
2-6	Earliest priority application: Application number	001
2-7	Earliest priority application: Filing date	2021-10-29
2-8de	Applicant name	Berthold R. Rutz
2-8	Applicant name: Name Latin	
2-9ae	Inventor name	dd
2-9	Inventor name: Name Latin	
2-10en	Invention title	fungal sequences
2-10fr	Invention title	c
2-10es	Invention title	f

图 123: 显示序列列表, HTML 样例

注: HTML 文件的位置将在用户浏览器的导航栏中显示, 以便用户可以将文件复制到其他位置。

当用户点击“Export Sequence Listing as .txt file”按钮时, 将打开 txt 文件。这提供了一个符合 ST.26 标准的 XML 文件的格式化视图, 使用户更容易看到特定字段的值。图 124 显示了一例。

```

Sequence Listing Information:
  DTD Version: V1_3
  File Name: validSTS.xml
  Software Name: WIPO Sequence
  Software Version: 1.1.0-beta.7
  Production Date: 2021-07-06
General Information:
  Current application / IP Office: US
  Current application / Application number: 1231123343
  Current application / Filing date: 2019-05-02
  Current application / Applicant file reference: app_file_ref
  Earliest priority application / IP Office: US
  Earliest priority application / Application number: 1231123343
  Earliest priority application / Filing date: 2019-04-30
  Applicant name: Vault Tec
  Applicant name / Language: en
  Inventor name: Vault Tec
  Inventor name / Language: en
  Invention title: FEV ( en )
  Invention title: fdf' ( ru )
  Sequence Total Quantity: 3
Sequences:
  Sequence Number (ID): 1
  Length: 368
  Molecule Type: DNA
  Features Location/Qualifiers:
    - source, 1..368
      > mol_type, other DNA
      > organism, synthetic construct
    - STS, 1
    - STS, 2..4
  Residues:
  atcatgctaa tcatgctagc tagtagctga tgatcatgct agcatcatgc taatcatgct 60
  agctagtagc tgatgatcat gctagctagt agctgatgat catgctagct agtagctgat 120
  gatcatgcta gctagtagct gatgatcatg cttagctagta gctgatgata atgctagcta 180
  gtagctgatg atcatgctag cttagtagctg atggctagta gctgatgtag tagctgatga 240
  tcatgctagc tagtagctga tgatcatgct agctagtagc tgatgatcat gctagctagt 300
  agctgatgat catgctagct agtagctgat gatcatgcta gctagtagct gatggctagt 360
  agctgatg                                     368

  Sequence Number (ID): 2
  Length: 368
  Molecule Type: RNA
  Features Location/Qualifiers:
    - source, 1..368
      > mol_type, genomic RNA
      > organism, Asaccus elisae
    - gene, 1
    - gene, 2..4

```

图 124: 显示序列表, TXT 样例

如果生成的 XML 序列表的大小超过 100Mb, 那么将显示图 125 中提供的 HTML 页面, 而不是显示 HTML 格式的序列表。



图 125: 显示序列列表，HTML 过大而无法显示

5 文件格式

以下文件格式可以导入到 WIPO Sequence。

ST.25

有关产权组织标准 ST.25 文件格式的详细信息，请参考：

<https://www.wipo.int/export/sites/www/standards/en/pdf/03-25-01.pdf>

MULTI-SEQUENCE (多序列)

多序列格式¹³可以描述一个或多个序列及其名称、分子类型和生物体名称。它是允许使用 PatentIn 导入的格式之一。

非空白文本的第一行是标头，由以下组件构成：

```
<SequenceName; SequenceType; OrganismName>
```

序列数据从标头之后的那一行开始。新序列在前一个序列的遗传密码结束之后，由文件中的新行划定。下面是以多序列格式定义的一组两个序列的示例。

示例：

```
<First Sequence; RNA; Albies alba>
```

```
uuuucuuauguuuucuccuacugcuuaucuaaauaugaugucguaguggcuuccucaucgucucuccccacc  
gccuaccacaacgacugccgcagcggauuacuaauaguaucaaccaacagcauaacaaaaagaugacgaa  
gagggguugcugauggugucgccgacggcguagcagaaggaguggcggagggg
```

```
<Second Sequence; DNA; Albies alba>
```

¹³ https://www.uspto.gov/sites/default/files/patents/resources/tools/checker/patentin351_20110214_6.pdf

```
attgaugtuagtgauguggtautgaugtuagutguagtautgaugtauuuuuuugtggtgagututuga  
uatguaautgautugtugutattgaugtuagtgauguggtautgaugtuagutguagtautgaugtauu  
uuuuugtggtgagututugauatguaautgautugtugutuagt
```

RAW

这种格式只能描述一个序列。遗传密码以其基本形式写入，没有附加信息。导入时，必须通过该工具将分子类型、特征和名称添加到序列中。

示例：

```
aggatatagatagtatatgatagtagtatgatgatgatgtatgtatagtagttatga
```

FASTA

该格式包含残基和说明，导入时用户可以选择将说明保存为 **note** 限定符。

有一个序列的 **FASTA** 文件

```
>AJ011880.1 Artificial oligonucleotide sequence SSR primer (CAC13R)  
CTCAACAATCTGAAGCATCG
```

参见 <https://www.ncbi.nlm.nih.gov/nuccore/3724029?report=fasta>（2017 年 5 月 22 日访问）。

[文件完]