WIPO
WORLD
INTELLECTUAL PROPERTY
ORGANIZATION

WIPO Sequence Version 2.0.0 User Manual

The purpose of this document is to provide the users with instructions on how to perform basic operations with the WIPO Sequence desktop application. Typically, the users are a patent applicant, or their representative, seeking to submit a patent application which includes a sequence listing.

Last Updated April 2022

WIPO SEQUENCE Version 2.0.0

USER MANUAL

TABLE OF CONTENTS

Page #

1 INTRODUCTION

1.1 OVERVIEW

WIPO Sequence is a desktop tool which enables a user to:

(i) create/edit a sequence listing in XML format compliant with WIPO Standard ST.26
(ii) verify the compliance of a sequence listing in XML format against WIPO ST.26 requirements.

The WIPO Standard ST.26 can be found at:

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

This document describes how to use WIPO Sequence as an applicant or a representative of an applicant. A full list of the functionality of the tool can be found in section 2 of this document.

1.2 SYSTEM REQUIREMENTS

The WIPO Sequence tool was developed to support the following Operating Systems:

- Windows 10 version 1803 (32- and 64-bits versions)
- Linux: Ubuntu version 18.04 and CentOS 7 version 1804
- MacOS version 10.13 (64 bits version)

Besides these versions above, it also supports the following operating systems:

- Windows 7 and higher (both 32bits and 64 bits)
- Ubuntu version 12.04 and newer
- MacOS version 10.9 (64 bits version)

The WIPO Sequence tool requires the following minimum hardware specifications:

- CPU: 1.6 GHz
- RAM: 4 Gb
- Free Hard Disk: 1 GB (additional HD can be required for storing the sequence listing information)
- Screen resolution: 1366x768

1.2.1 Installation

1.2.1.1 Windows

WIPO Sequence provides a single installation file for both 32- and 64-bits versions of Windows. The user should follow the process shown in the installation wizard.

We have to specify that users will have performance problems, or the tool will not work properly if the 32bit version is used.

The database files used for storing the project information along with the log of the tool are stored at the following location:

```
C:\Users\<user>\AppData\Roaming\ST26_authoring
```

When the application is updated or uninstalled, these files are not removed, so the projects data will remain if the application is reinstalled.

1.2.1.2 Linux

WIPO Sequence is provided as an “AppImage” file (https://appimage.org/) that will run on most Linux distributions, including CentOS and Ubuntu. In order to run the file, the user can execute the file by double-clicking on it or executing using the command line.

Initially, the user will be prompted with a message to select whether a shortcut on the desktop should be created.

1.2.1.3 OSX

WIPO Sequence provides a “dmg” file for installing the application on a MacOS 64-bit operating system. In order to install it, the user should double-click on the file and follow the wizard.

The database files used for storing the project information along with the log of the tool are stored at the following location:

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

When the application is updated or uninstalled, these files are not removed, so the projects data will remain if the application is reinstalled.

1.2.1.4 Silent install

WIPO Sequence supports a silent install by use of the following flags during installation (with .exe installer file):

- */S*: to launch a silent install
- */allusers*: to installs the tool so it is available for all Windows users on the desktop machine (this has to be launched when logged in as an admin user).

1.2.2 Uninstall

1.2.2.1 Windows

WIPO Sequence provides an uninstall wizard that can be launched under the “Add or Remove Programs” option in Windows.

In order to completely remove the log files and the files used to store the projects information, the following folder must be deleted:

```
C:\Users\<user>\AppData\Roaming\ST26_authoring
```

1.2.2.2 Linux

By removing the Linux “AppImage” file, the application is uninstalled from the computer. Additionally, the menu entry can be removed by deleting the desktop file from the location:

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

In order to completely remove the log files and the files used to store the projects information, the following folder must be deleted:

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

1.2.2.3 OSX

The application can be uninstalled from Finder on OSX, under the Applications section.

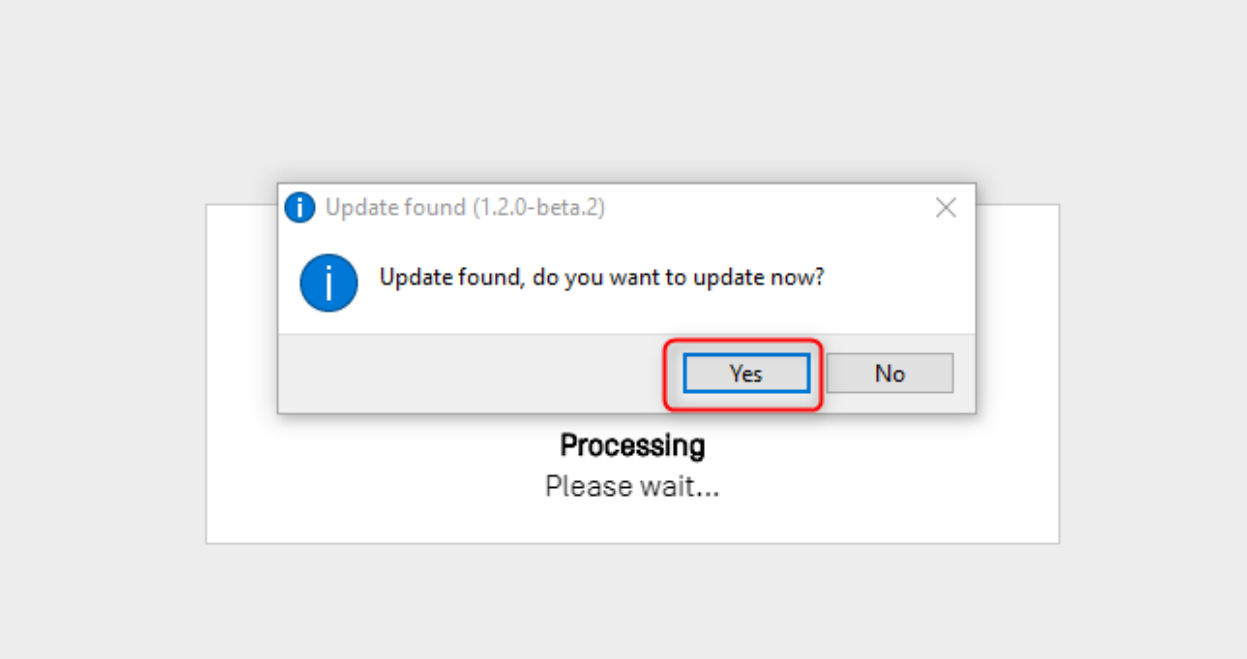
In order to completely remove the log files and the files used to store the projects information, the following folder must be deleted:

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

1.2.3 Auto update process

After launching WIPO Sequence, the tool will check for available updates and if an update is found, it will ask the user to upgrade to the newer stable version.

It is advised to not skip the update to ensure the version used conforms to the latest version of the ST.26.



Note:

In order to enable the auto update, the computer must be connected to the internet and the tool must have permissions to access the WIPO website through HTTP. Please also note that no information is sent from the user to the WIPO update server.

The user should wait patiently for the new version to download and install rather than change focus from the application.

1.2.1 Log4J Vulnerability

A zero-day exploit has been reported against the popular Log4J2 Javascript library which can allow an attacker to remotely execute code[1].

According to the Spring Boot documentation, only applications using log4j-core and including input in log messages are vulnerable and as the WIPO sequence validator uses only the log4j-to-slf4j and the log4j-api and both cannot be exploited on their own. As such, the WIPO Sequence is not affected by this vulnerability.

2 TOOL FUNCTIONALITIES

This section outlines all the functionalities that are implemented by the tool, in the current version of the tool, version 2.0.0. The list of functionalities is provided here in alphabetical order.

ID	Functionality
1	Add a custom organism name & description to the list of organism names of the system
2	Add an invention title and its corresponding language code to a project
3	Add application information (either current or prior application) to a project
4	Add feature information to a sequence
5	Add a source feature and its mandatory qualifiers to a sequence
6	Add new organism names to the list of organism names stored in this system
7	Add qualifier information to a feature

[1] https://spring.io/blog/2021/12/10/log4j2-vulnerability-and-spring-boot

8	Add sequence listing general information data to a project
9	Add a sequence name to a selected sequence
10	Change the order in which the sequences will be listed in the generated sequence listing
11	Create a workspace in which data related to one sequence listing is stored
12	Create an instance of the sequence data structure and set its attributes to the values obtained from a ST.26 sequence data XML node received as input
13	Create and insert a sequence in another position in the listing
14	Create a translation qualifier for a selected CDS feature and its associated translated sequence
15	Create a new person or organization name
16	Delete a sequence
17	Display the generated sequence listing
18	Display the data changed upon import of a sequence
19	Edit the attributes of an instance of “Qualifier” data structure
20	Edit the attributes of a project
21	Edit the attributes of a sequence
22	Edit the attributes of an instance of “Feature” data structure
23	Enable or disable selected verification rules
24	Export of free text qualifiers, for the purpose of translation in XLIFF format
25	Export all data stored in a project so that it can be later imported into the same or a different instance of the system
26	Export the list of custom organism names to a file that can be later on imported into a different instance of this system

27	Export Generated sequence listing in human readable format (.html & .txt)
28	Generate sequence listing
29	Import a list of custom organism names from a file
30	Import all data stored in a project file
31	Import data from a ST.25 sequence listing file into a newly created project
32	Import data from a ST.26 sequence listing file into a newly created project
33	Import data from a FASTA file into an existing project
34	Import multiple sequences from a file in “one go”
35	Import into the current project (target project) the data from another project (origin project)
36	Parse a multi-sequence format sequence and review it returns the four parts (sequence name, molecule type, organism, and residues)
37	Print data from the project or generated ST.26 sequence listing
38	Provide location information to a selected feature
39	Provide the system with a string of residues containing invalid symbols and verify the process residues reformatted
40	Record data that has been changed upon import, so that the original data and the changed data can be displayed after import
41	Remove from the project all data associated with a sequence and renumber the remaining sequences accordingly
42	Set the “INSDQualifierMolType” attribute of a sequence to one of the predefined values
43	Set and store the system preferences (maximum number of residues symbols displayed per line, etc.)
44	Store in the system information about an applicant or inventor, (for example, name, its corresponding language code and its translation or

	transliteration into Latin characters (if applicable), address etc.) so that it can be used later in various projects
45	Translate a nucleic acids sequence according to a specified genetic code table number
46	Verify a ST.26 sequence listing file and list the issues as a verification report containing warning and error messages
47	Verify the data stored in a project and list the issues as a verification report containing warning and error messages
48	Import of free text qualifier source-target pairs in XLIFF format, for the purpose of providing translations in a project
49	Bulk editing of sequence annotation including qualifier mol_type

3 TOOL OVERVIEW

3.1 TOOL MAIN ELEMENTS

This section details the main elements of the tool. The goal in providing this detail is for the user to be familiar with the common components provided within the desktop tool.

3.1.1 Page

The Page is the main container for views within the tool. There are two main pages that contain views within them:

- Main Page
    - Projects Home View
    - Persons and Organizations View
    - Custom Organisms View
    - Preferences View
- Project Page (accessible from the Projects View)
    - Project Detail View
    - Verification Report View
    - Language Dependent Qualifiers View
    - Import Report View
    - Display Sequence Listing View

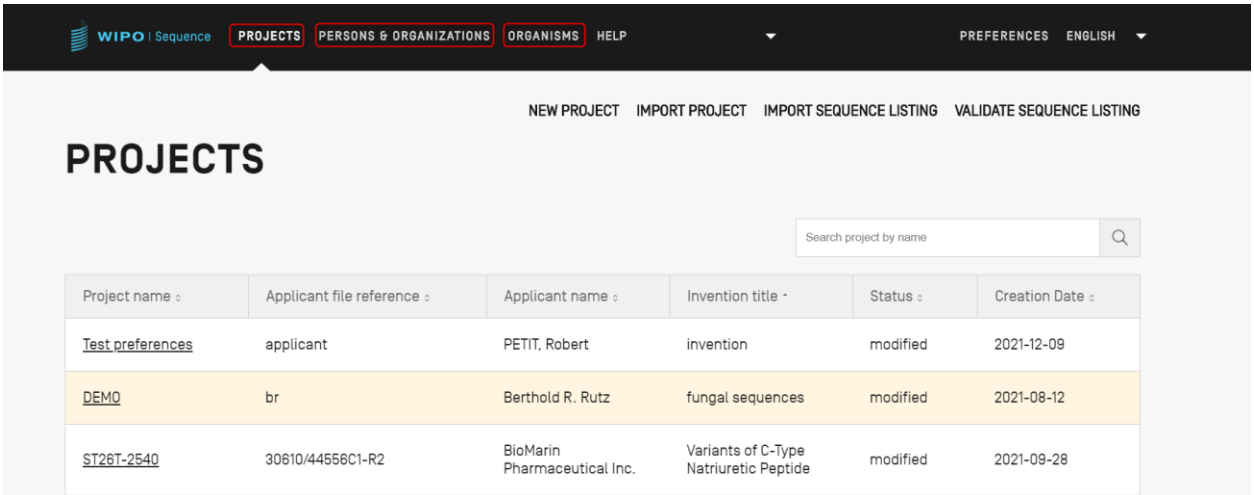


Figure 1 Main Page

3.1.2 View

The different displays of information that can be seen within a same page are termed Views.

3.1.3 Sections

Some views can have Sections. Sections provide a convenient way of compartmentalizing different parts of a large View.

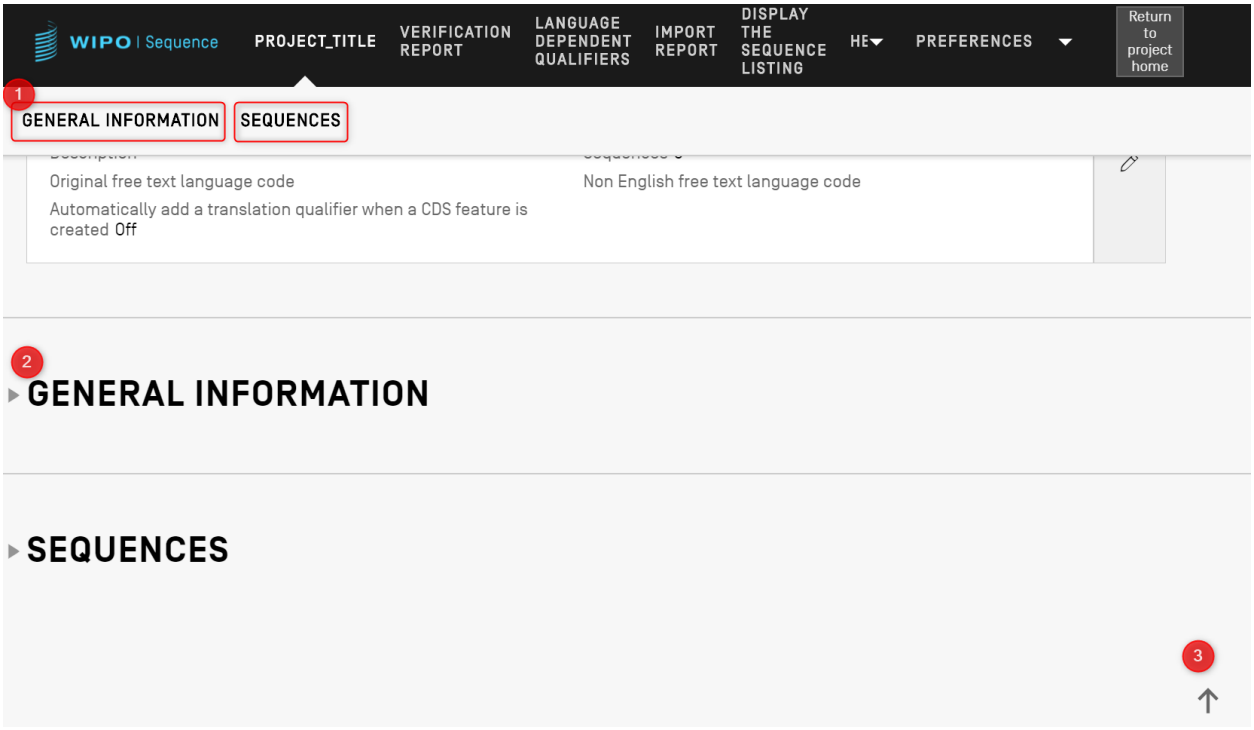


Figure 2 Project Detail view

Within the Project Detail View, shown in Figure 2, there are 2 collapsible sections: General Information & Sequences. At the top of the Project Detail View, there are two links that will navigate the user to the corresponding sections (1); each section is collapsible (indicated by the small triangle to the left) for the sake of ease of navigation (2); the user can click on the arrow icon to scroll to the top of the Project Detail view (3).

3.1.4 Overlay

When a Panel needs to be filled out or modified, sometimes an Overlay will appear over the current view, greying out the background as shown in the following example captured in Figure 3.

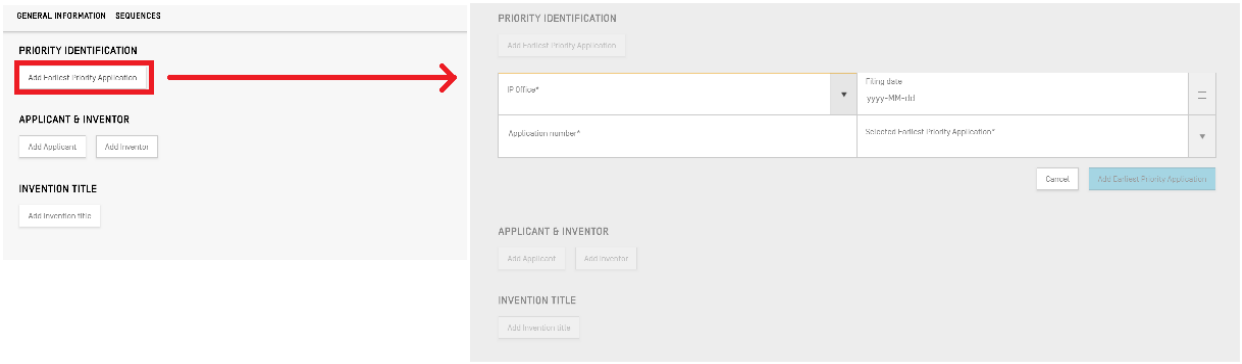


Figure 3 Overlay

When the Overlay is visible, all elements behind the Overlay are non-functional, and only the items in focus will be modifiable.

3.1.5 Tables

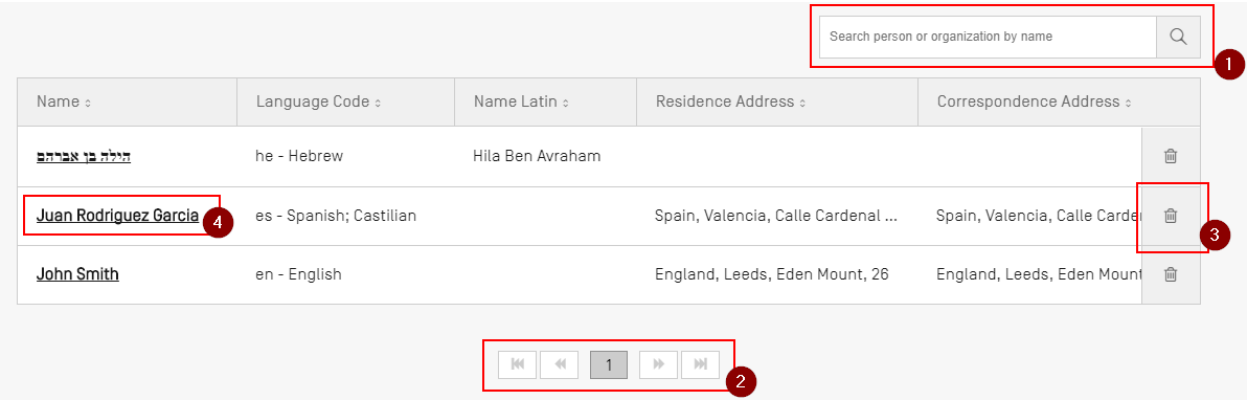


Figure 4 Tables

An example of a Table is shown in Figure 4. When search is enabled for the Table, the user can enter some search data and click on the search icon to the right of the search bar (1). To return to the full list of elements within the table, clear the search input box and click on the search icon (1).

The user can navigate through the pages of rows in a Table by clicking on the buttons below the table, in the case that all the entries do not fit in one page (2). To delete an entry in a Table, the

user can click on the trash icon to the right of the corresponding row (3). To open the edit view of an entry in a Table, the user must click on the underlined value in the row (underlining indicates it is clickable) (4).

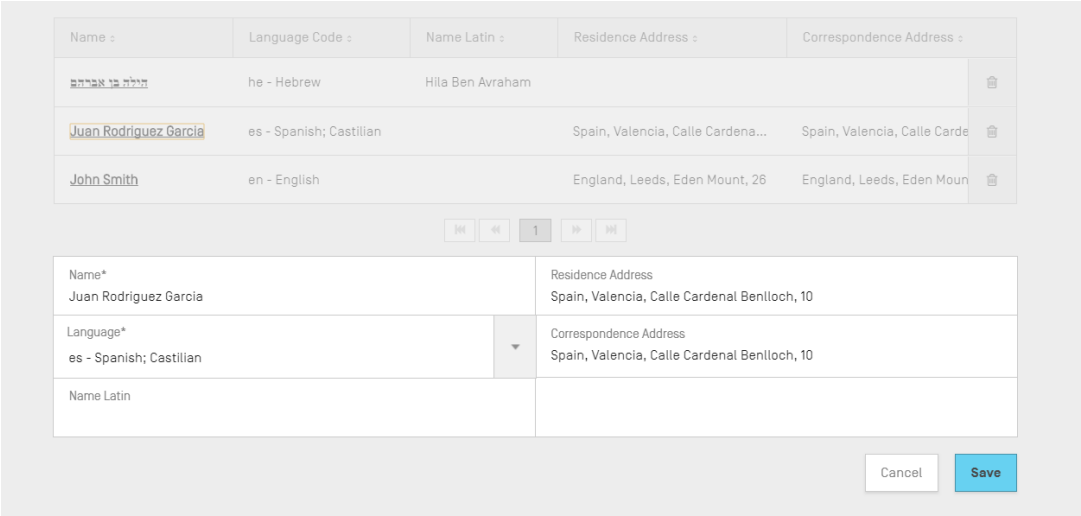


Figure 5 Table edit

Once this entry has been selected, an Overlay will appear over the screen and an editable panel will open beneath the Table.

3.1.6 Panel Views

Panel Views display a group of data. Items in a Panel View can be distributed across several columns. Each item has a label and a (optional) value.

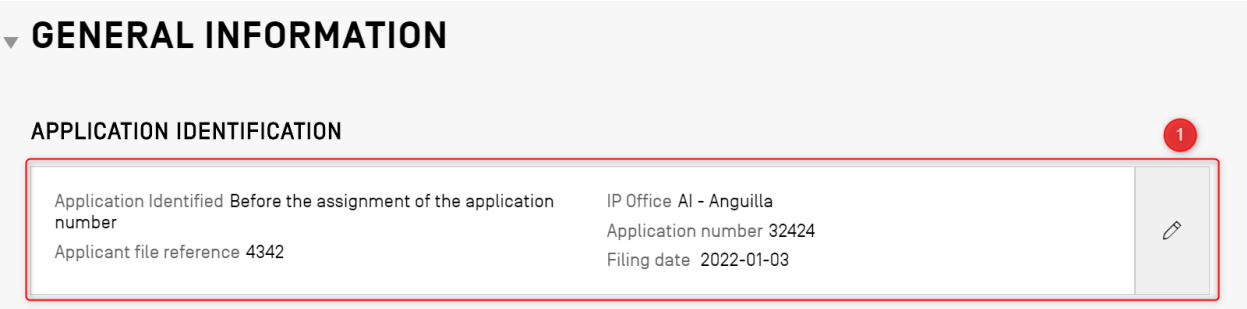


Figure 6 Panel View

As shown in Figure 6, when the button with the pencil icon (1) is clicked on a panel view, the Panel View is replaced with an Edit Panel.

3.1.7 Edit Panel

Once an Edit Panel has been opened, the fields that can be modified are presented to the user. Once the user has finished editing values, they will be able to either save the changes made, or discard them by clicking on the “Save” (1) or “Cancel” (2) buttons respectively (shown in Figure 7).

Note:

Fields that are only for display, and not modifiable, will appear greyed out.

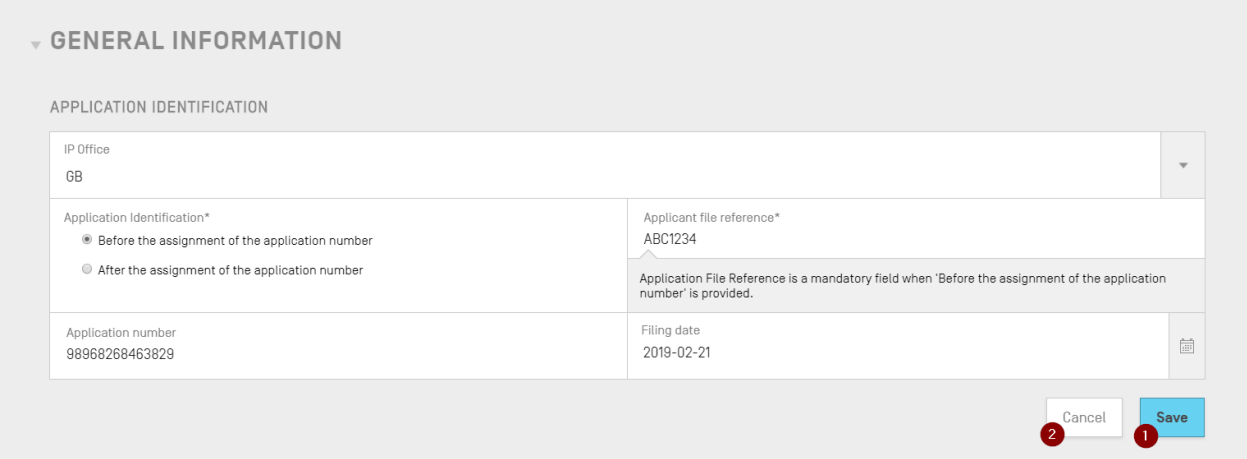


Figure 7 Edit Panel

3.1.8 Date Picker

Whenever a user is required to enter a date, they will use the Date Picker interface. By clicking on the calendar icon (1), the Date Picker component will open (2), as shown in Figure 8.

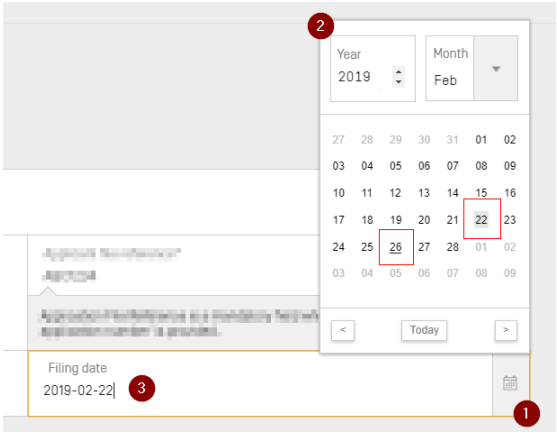


Figure 8 Date picker

The currently selected date is marked with a grey background (22nd), and the day of use is indicated with a black underline (26th). The first day of the week is considered to be a Sunday.

Conversely, the user can also simply type in the desired date in the appropriate format ("*YYYY-MM-DD"*) (3).

3.2 PDF READER

When a user prints a Verification report for a particular sequence listing, an Import report or Project information, a PDF file will be generated and opened in a PDF Reader. To download the file and save it, the user must click on the download icon at the top right of the viewer (1), shown in Figure 9.

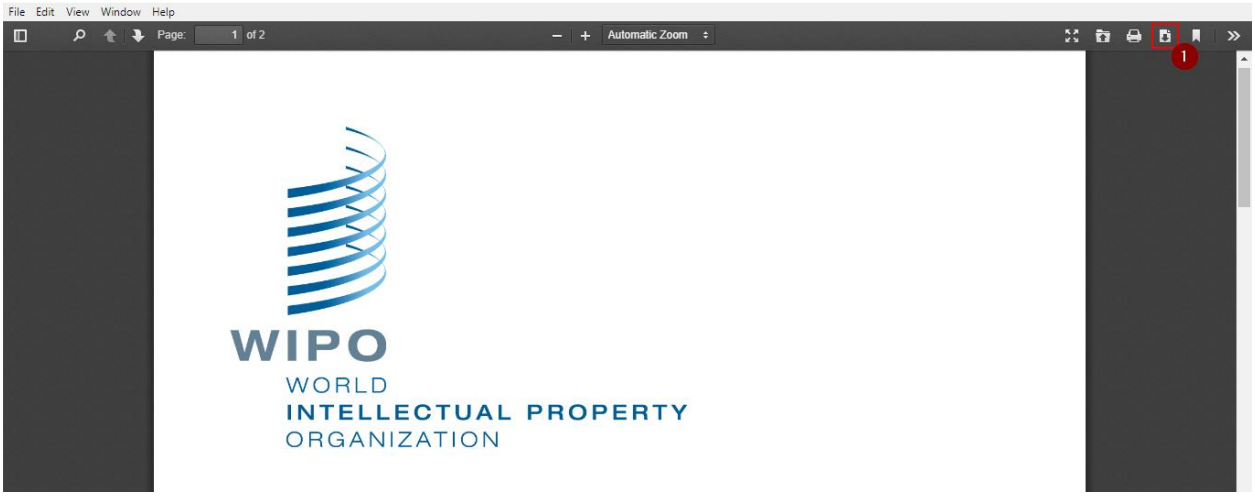


Figure 9 PDF Viewer

3.3 KEYBOARD NAVIGATION

WIPO Sequence supports basic keyboard navigation. The 'TAB' key is used to navigate between items and the 'SPACE' key is used to select checkbox and radio buttons.

The focus during the navigation is visible for Edit button:

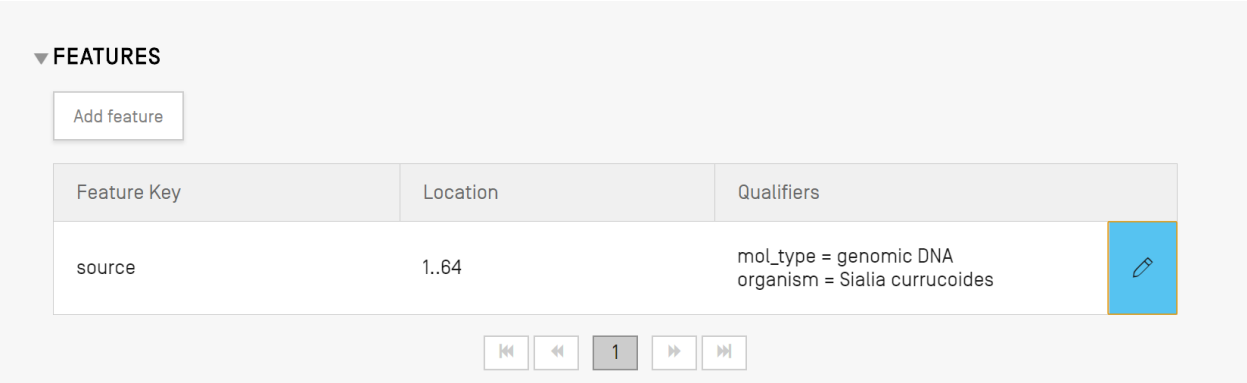


Figure 10 Focus on Edit Feature Button

Also, setting the cursor and the focus to the top left input field.

For example, create new project, there is only one mandatory field, the cursor and focus already set to project name field (Figure 11) bellow:

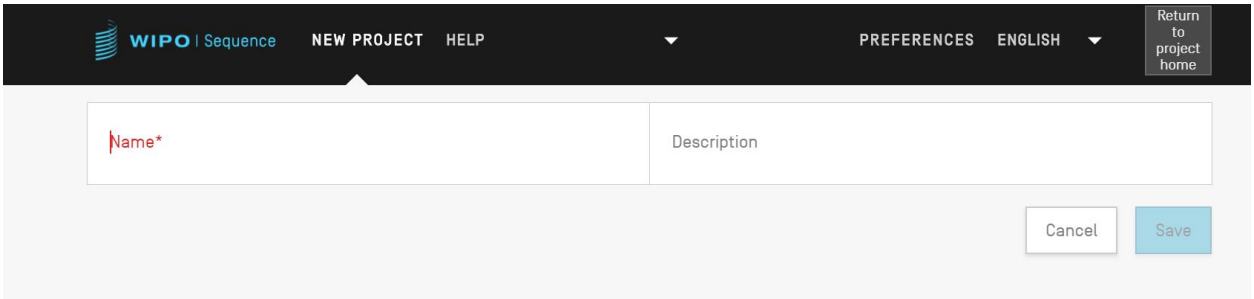


Figure 11 Cursor focus on the mandatory field

4 MAIN PAGE

The main page of the tool consists of 3 main Views:

The ***Project*** View (Figure 12), the ***Persons & Organizations*** View (Figure 13) and the ***Organisms*** View (Figure 14).

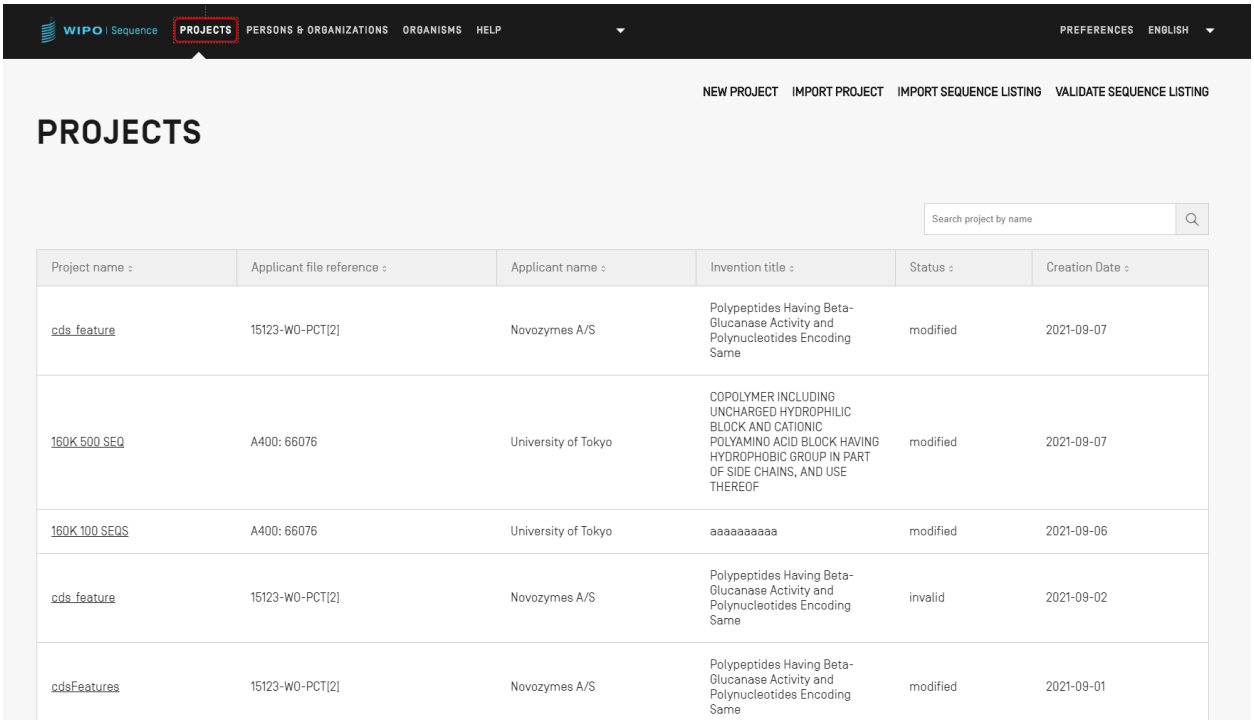


Figure 12 Projects View

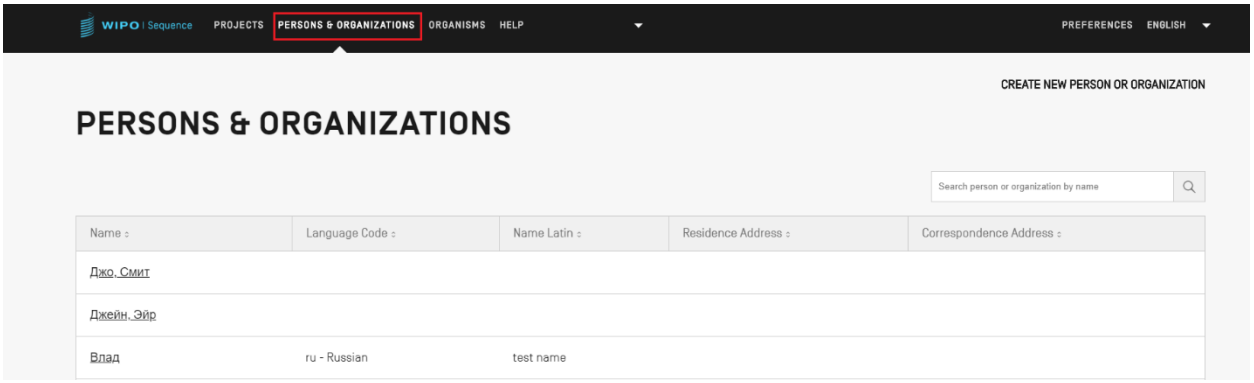


Figure 13 Persons & Organizations page

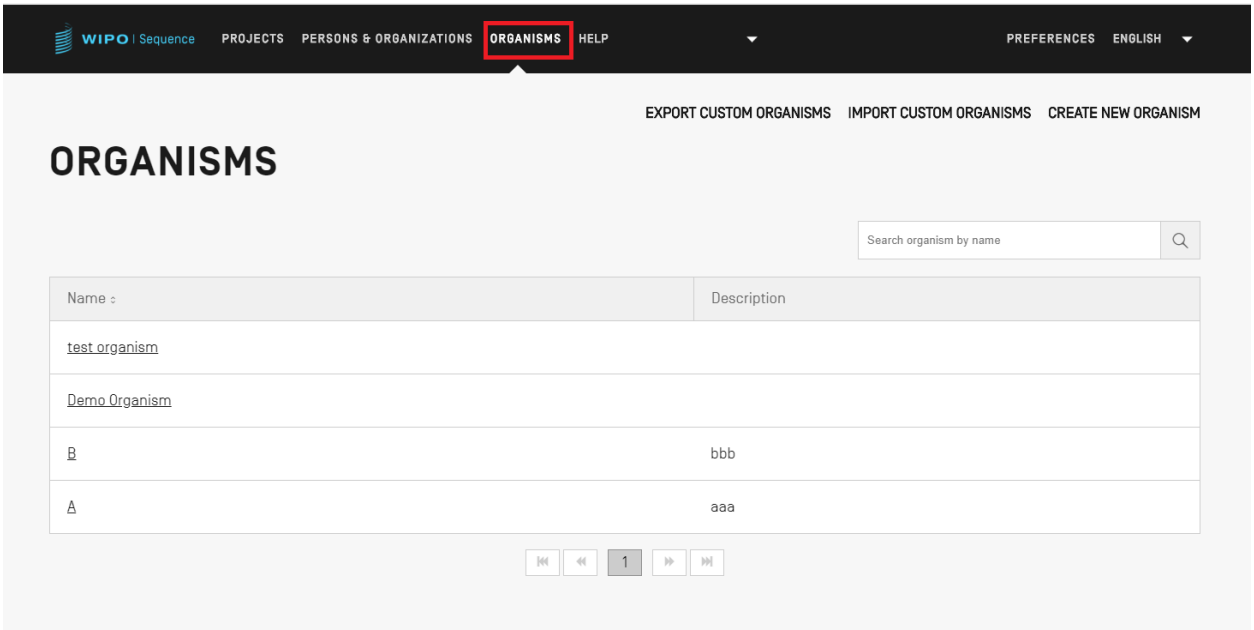


Figure 14 Organisms view

4.1 PROJECTS HOME VIEW

This section details the different options accessible in the Projects Home View.

A project is the object structure that the tool uses to store data necessary to generate a sequence listing. The tool uses data stored in the project, once this data has been validated as compliant with WIPO Standard ST.26, as the values within the generated sequence listing.

On this view the list of the created projects is displayed, giving the user the option to sort or use the search function to filter by project name, applicant file reference, applicant name, invention title, status or creation date.

Note:

The tool displays a max of 1000 projects. If a project is not displayed in the Projects Home view, the user should use the search function to identify the project by its name as it will still be stored locally but just not visible in this view.

4.1.1 Create Project

To create a new project, the user must begin from the main Projects Home View (see Figure 15).

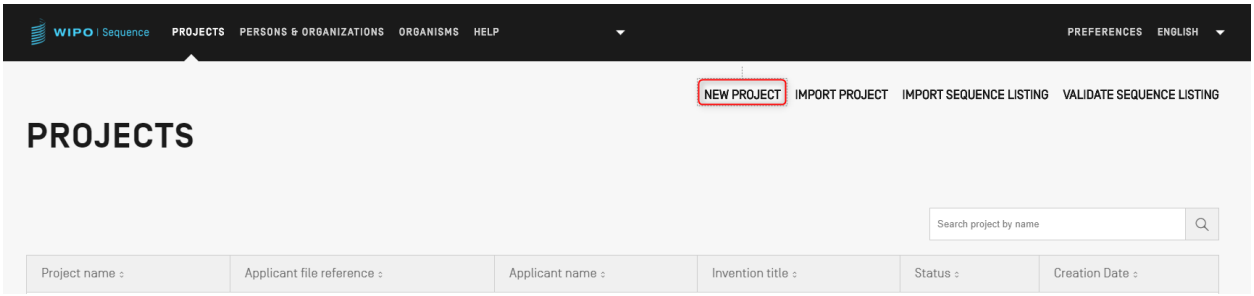


Figure 15 New project (a)

1) Click on the “NEW PROJECT” link at the top of the View indicated in Figure 15. In the following view, the tool will request a Name (mandatory) and a Description (optional).

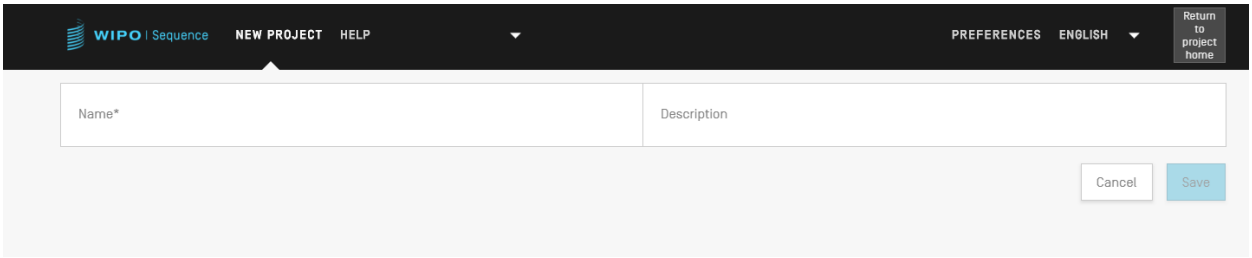


Figure 16 New project (b)

2) When a value is entered in the name field, the “Save” button will be enabled for the user to save the new project. The list of projects which includes this new project in the Project Home View is shown in Figure 17.

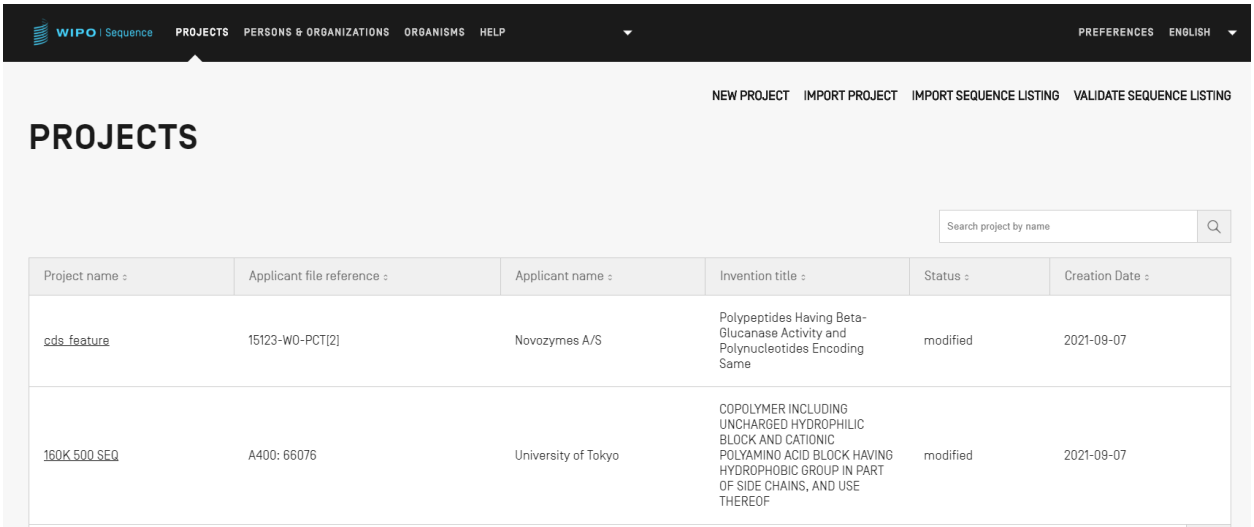


Figure 17 New project (c)

4.1.2 Import Project

This functionality allows the import into the tool of a previously exported project as described in section 5.1.1.4 Export Project. To import a project file, the user must begin from the Projects Home View.

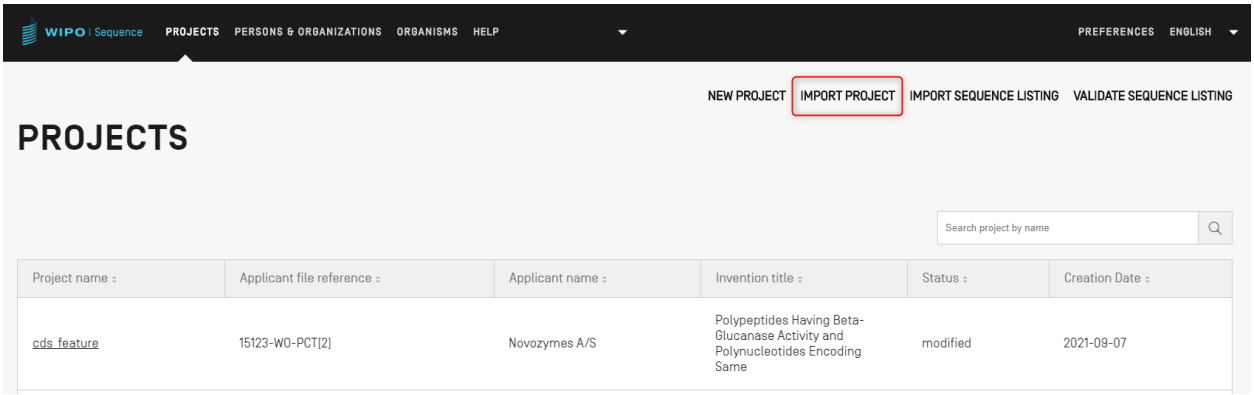


Figure 18 Import project

1) Click on the “IMPORT PROJECT” link at the top of the view as indicated in Figure 19.

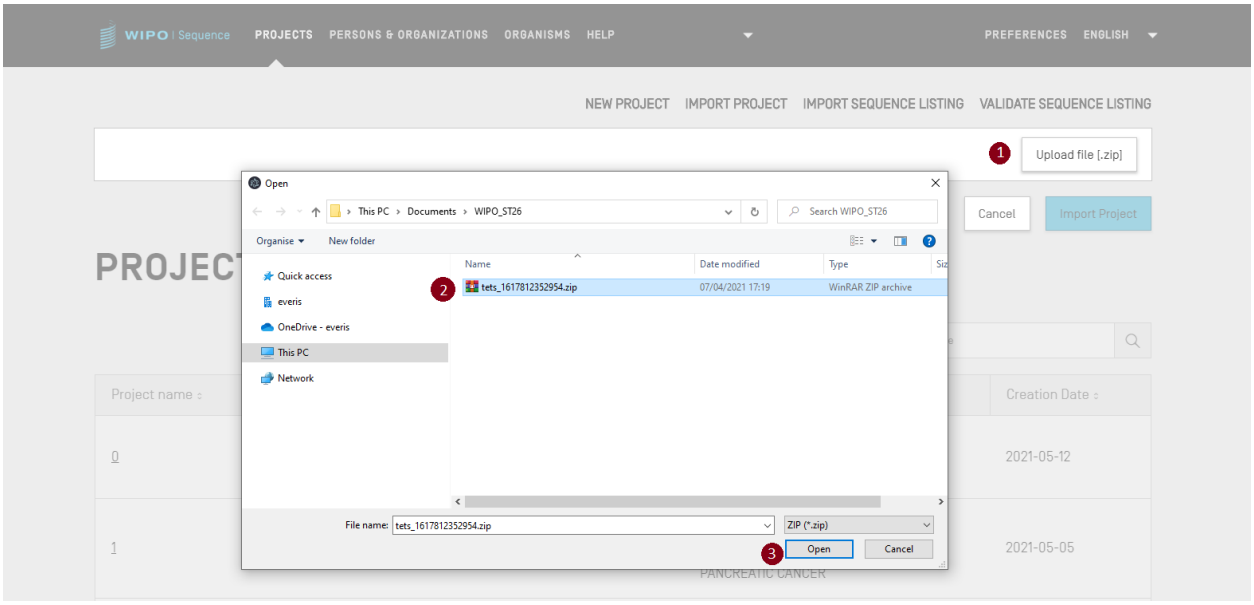


Figure 19 Import project dialog box

2) On the Overlay screen shown in Figure 19, click the “Upload file [.zip]” button (1).

3) In the dialog box that opens, shown in Figure 20, select the project file to be imported (2 & 3).

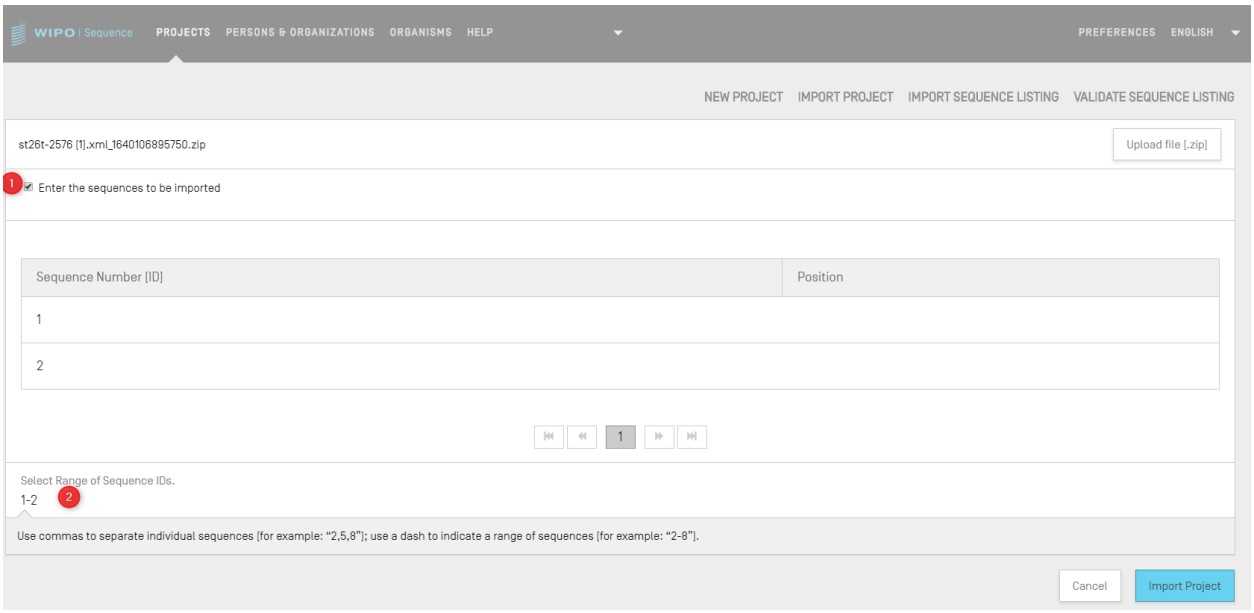


Figure 20 Import project sequence select

4) If “Select Range Sequences” remains unchecked, all the sequences will be imported. If the user wishes to select which sequences to import into the project, check the “Select Range Sequences” checkbox (1) and enter the ID numbers of the desired sequences in the appropriate field (2) shown in Figure 21. A single sequence can be entered, as well as a list of sequences separated by commas or a range of sequences in the form `x-y`.

   Example: “1, 3, 7, 13-20, 30-50”

   By default, the total number of sequences of the imported project will be displayed as a range i.e.,: 1- total sequences

5) The final step is to click on the blue “Import Project” button (3), as shown in Figure 21.

If the project is successfully imported, the following blue banner and message will appear at the top of the View.

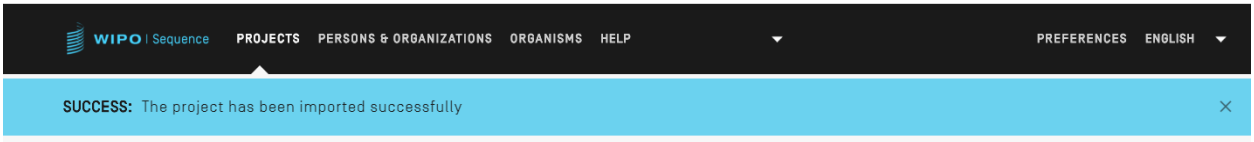


Figure 21 Import project banner

Known issue: WIPO Sequence can generate a lot of unexpected errors while importing the project: it is not clear what the cause of this error is. This mostly occurs when tool that is hanging. If you are having troubles, please try the import process again.

4.1.3 Import Sequence Listing

From the Projects Home View, the user can import exclusively the sequence information from a ST.26 ***or*** ST.25 compliant sequence listing. The file formats for each are *.xml for ST.26 format and *.txt for ST.25 files.

1) First, click the “IMPORT SEQUENCE LISTING” button at the top of the view shown in Figure 22.

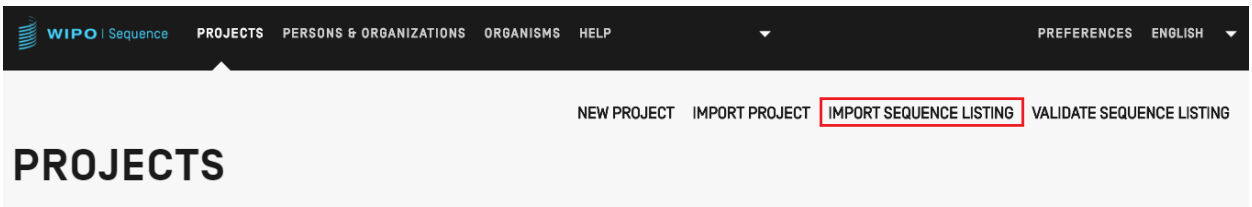


Figure 22 Import sequence listing

2) In the Overlay that opens, click on the “Upload file ST.25 [.txt] or ST.26 [.xml]” button (1), and select the desired sequence listing file in the dialog box that opens (2). Then enter a

name for the new project that is being created under the name given in the “Project Name” field (3).

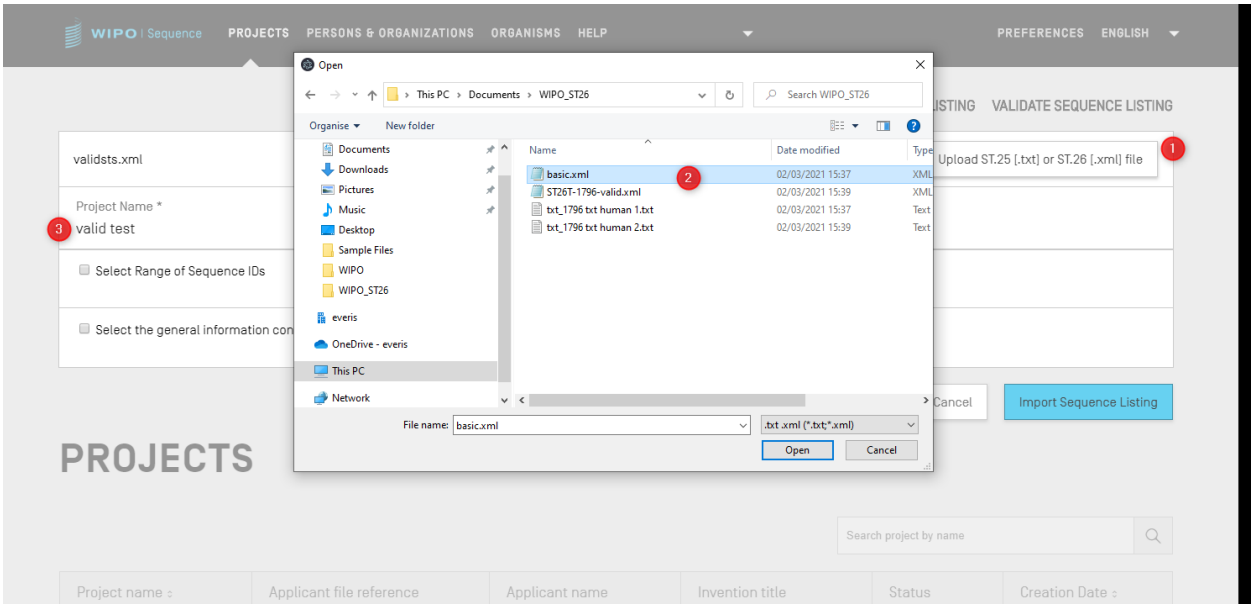


Figure 23 Import sequence listing (dialog box)

Two primary checkboxes shown in Figure 24 will allow the user to specify which sections are to be imported into the new project, “Select Range Sequences” and “Select the general information contents to be imported”.

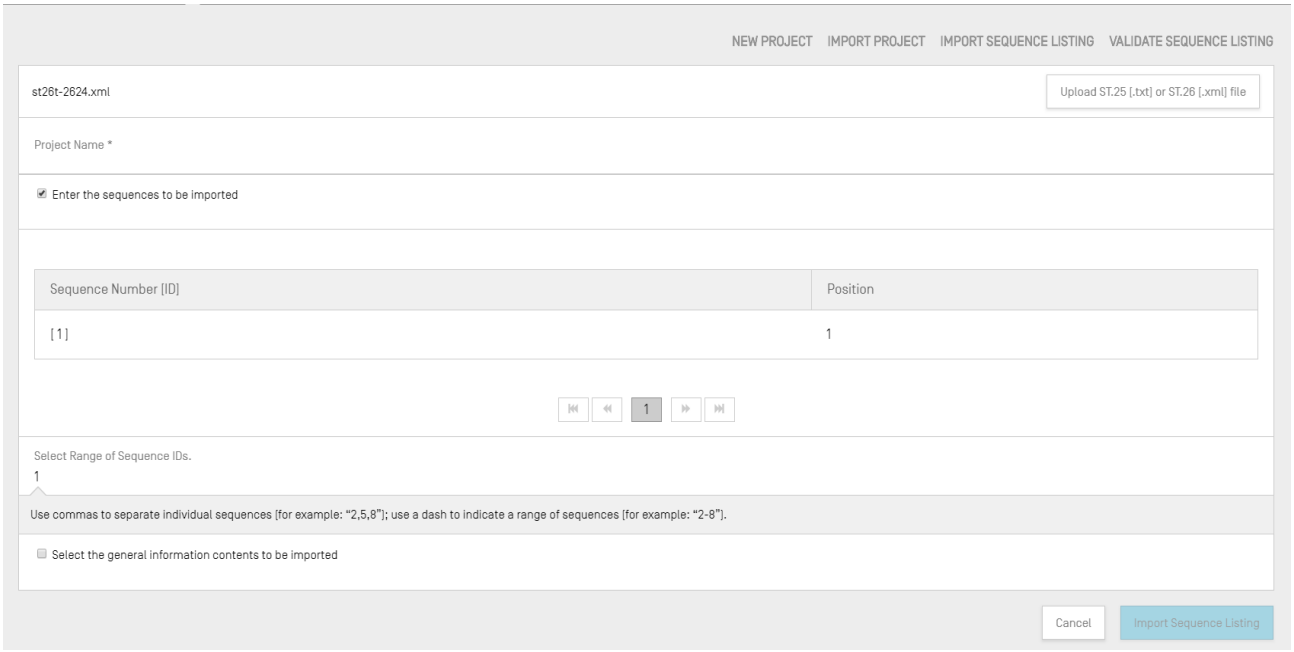


Figure 24 Import sequence listing, select range

The first checkbox will allow the user to enter which specific sequences they wish to import from the sequence listing. A single sequence can be entered, as well as a list of sequences separated by commas or a range of sequences in the form `x-y`.

Example: “1, 3, 7, 13-20, 30-50”

By default, the total number of sequences of the imported Sequence Listing will be displayed as a range.

The Table shown in Figure 24 has two columns, one gives the Sequence ID Number to identify the corresponding sequence and the other gives the “position” in which it will appear in the sequence listing.

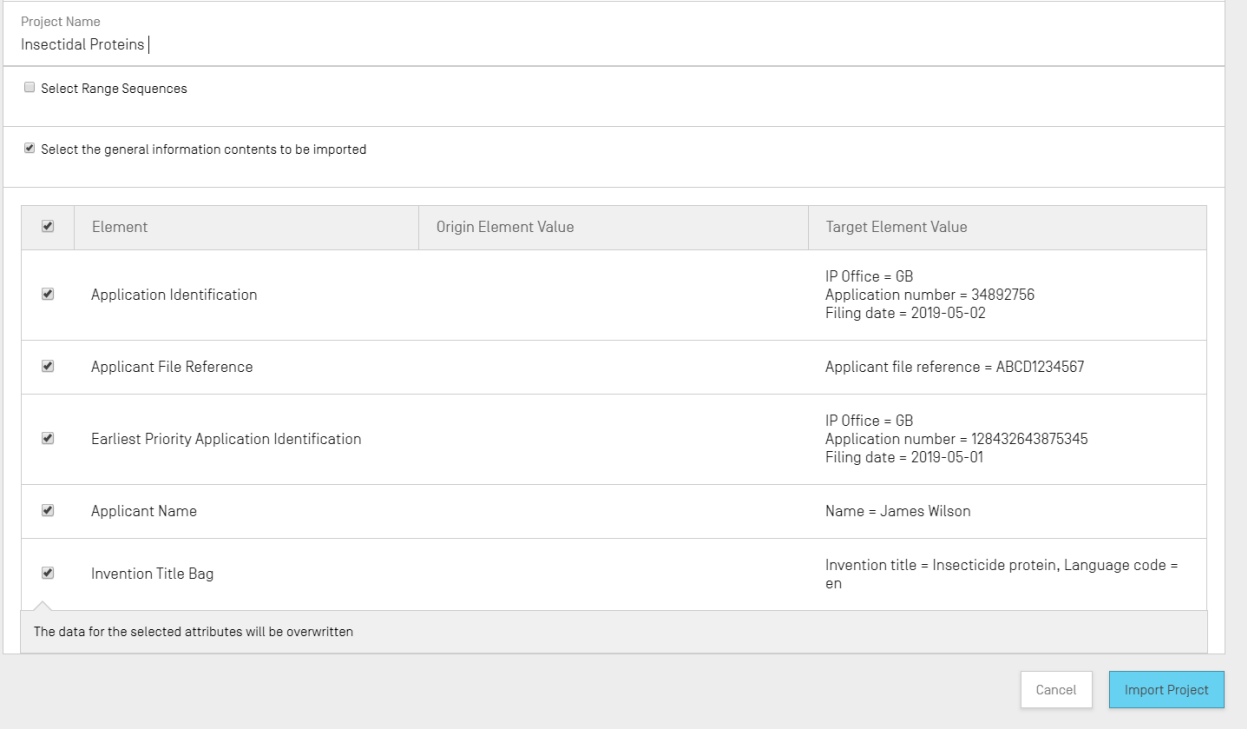


Figure 25 Import Sequence Listing, select General Information

Checking the second checkbox will enable a list of additional checkboxes which allow the user to individually select which properties are to be imported or ignored, as shown above in Figure 25.

3) Finally, click on the blue “Import Project” button to create the new project.

If the Sequence listing was imported correctly, the Change Data Table shown in Figure 26 will be displayed informing the user of the automatic changes made to the ST.25, multi-sequence and raw sequence listing data during import, in order to adapt it to the requirements of ST.26.

Note: that for importing a Sequence Listing the Features and Qualfiers are case sensitive and should comply with the values provided in Annex I of WIPO ST.26.

It is also important to note that ST.25 compliant sequence listings imported into must be valid as otherwise the functionality of WIPO Sequence cannot be guaranteed.

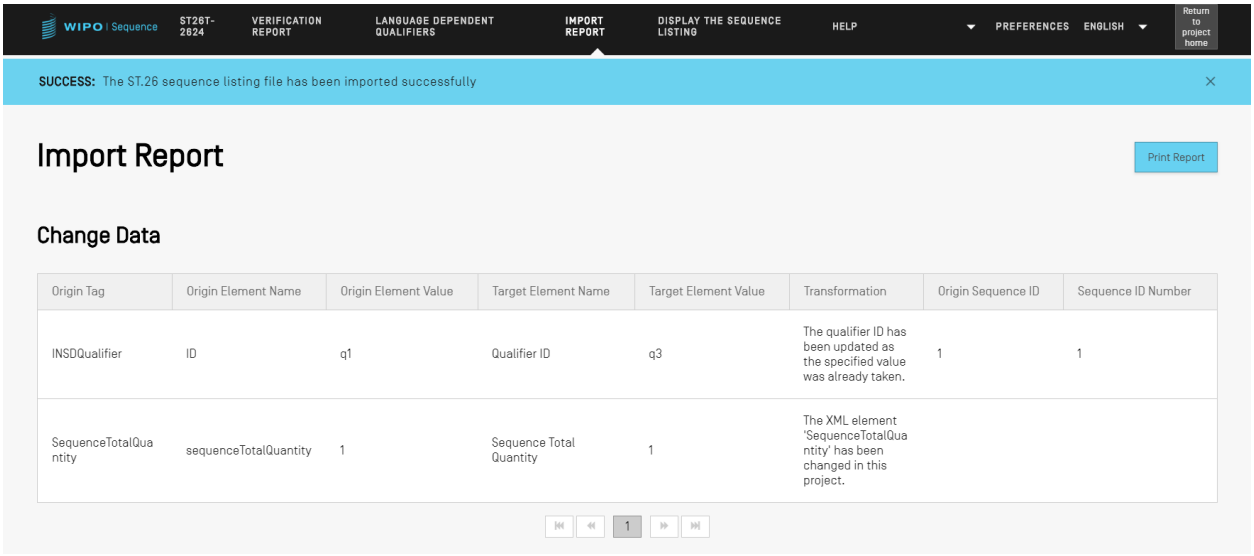


Figure 26 Import sequence listing, Change data

If the file format was ST.25, then the Import Report View will include an Import Report Table first, as well as the Change Data Table. An example of the Change Data report is shown in Figure 26 and an example Import Report is shown in Figure 27.

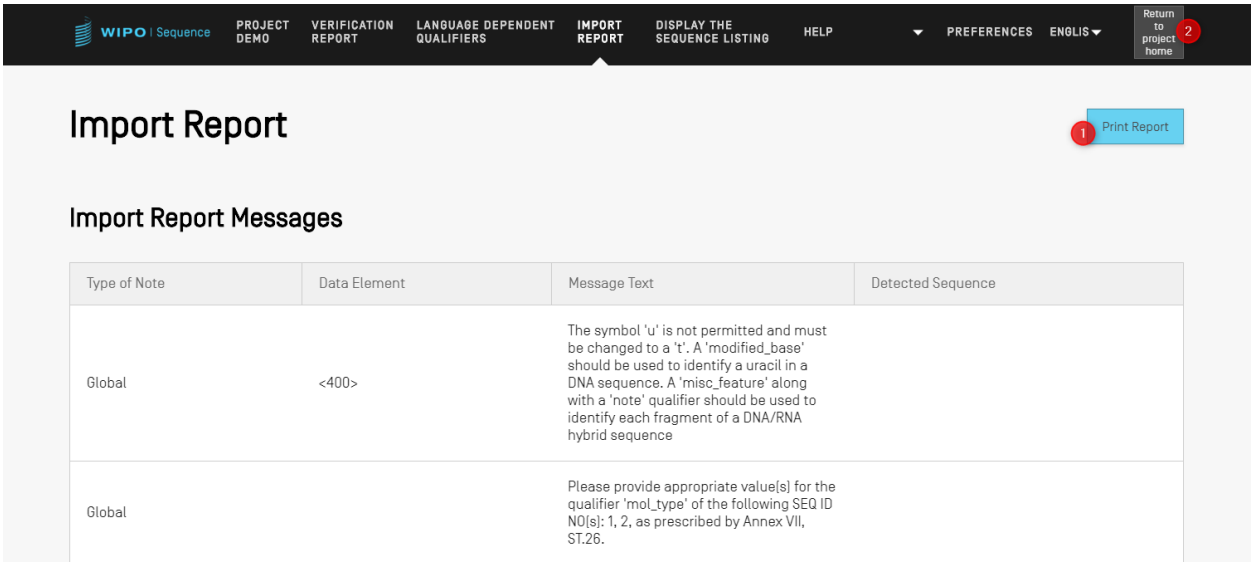


Figure 27 Import sequence listing, Import report

At this point, the user can return to the Projects Home View (2) or print a report of these changes in PDF format (1): see Figure 27.

For instructions on how to download the PDF file, see section **2.2.**

Conversely, the import process can fail if there are errors in the sequence listing file. In this case, after attempting to import, the user will be notified with a red banner indicating an error has occurred during import, see Figure 28.

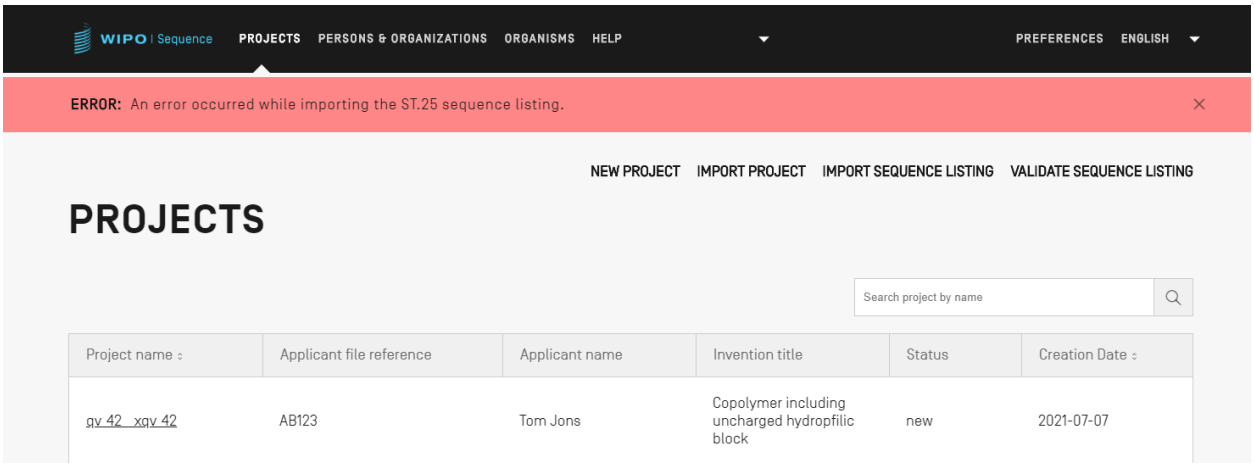


Figure 28 Import ST.26 sequence listing, Import report

In addition, the tool performs best at the threshold limit of 100k sequences. When dealing with large sequence listings, the user can perform the following workaround: split the import process into a series of steps by choosing a specific range of the sequences to import and then importing these sequence inside a project range-by-range. (For example, a sequence listing of ~100k sequences can be split into a series of 10 x 10k sequences and these can be imported one-by-one. The first 10k would be used in the creation of the project).

4.1.4 Validate Sequence Listing

The user can validate an ST.26 sequence listing file by clicking on the “VALIDATE SEQUENCE LISTING” button at the top right of the Projects View, shown in Figure 29.

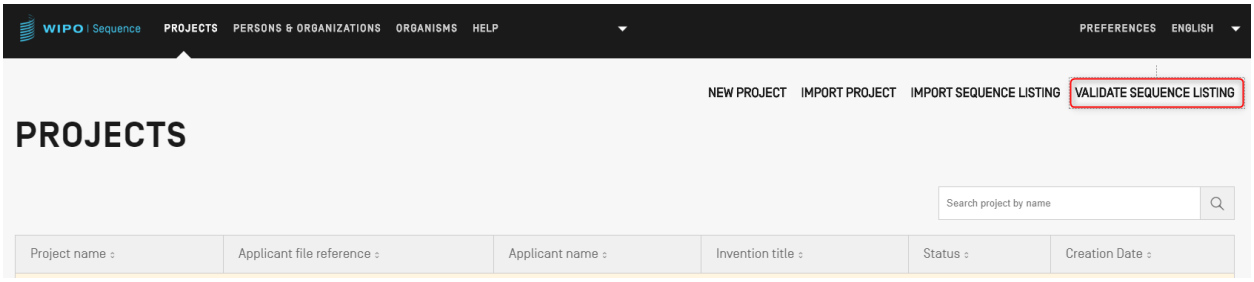


Figure 29 Validate sequence listing

Next, the user must click on the “Upload file ST.26 [.xml]” button (1), and then select the file in the dialog box (2).

Lastly, click the “Validate Sequence Listing” button (3).

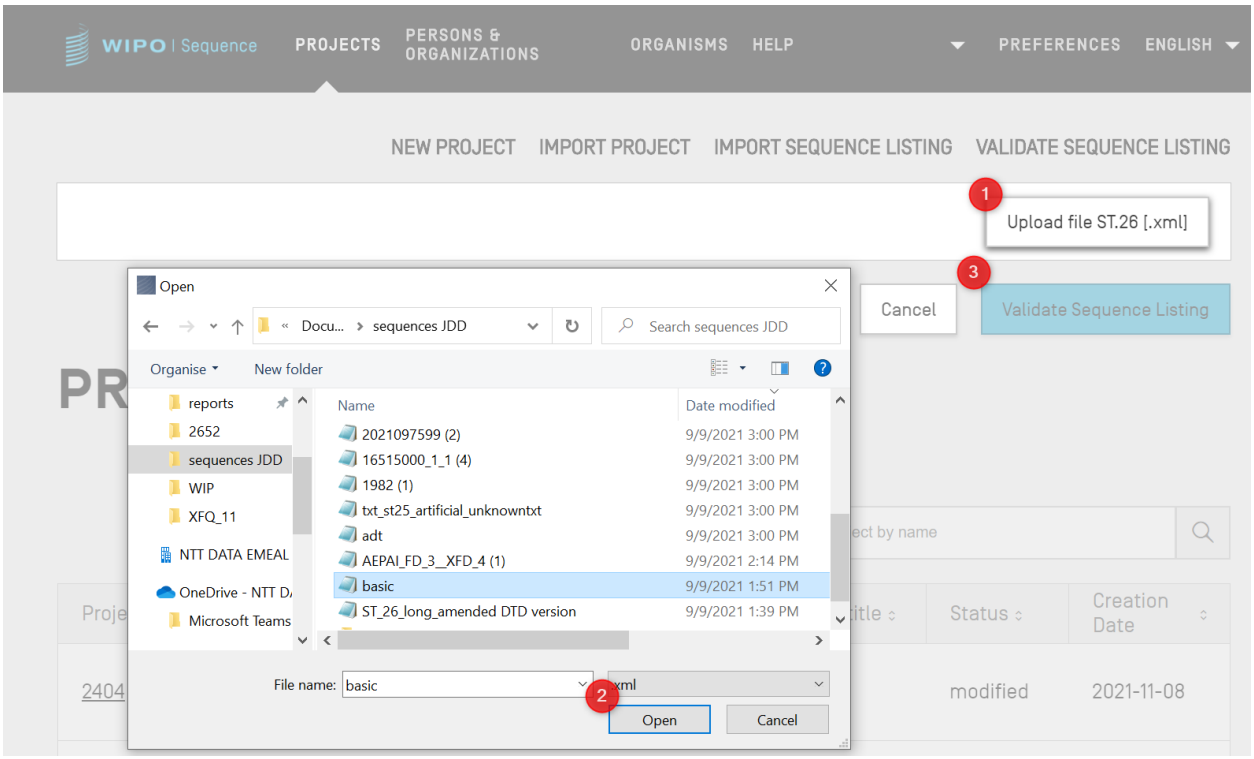
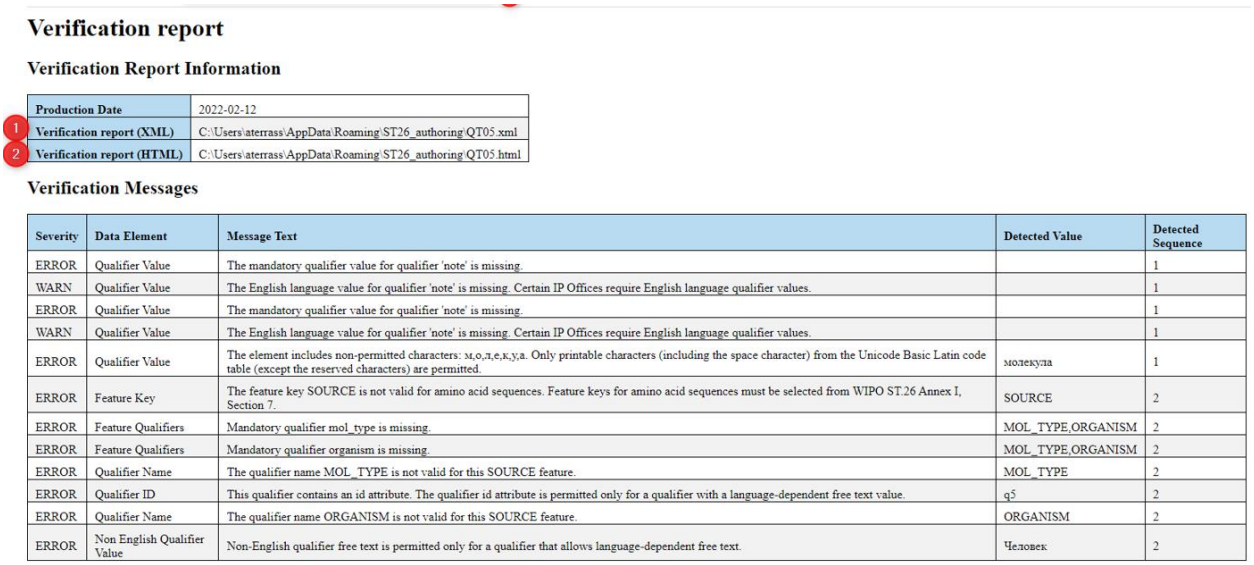


Figure 30 Validate sequence listing (dialog box)

If the sequence listing passes validation, a banner will appear as shown in Figure 31:

Figure 31 Validate sequence listing, Success banner

If the sequence listing fails the validation, a verification report will be opened in the user's browser with the validation errors listed in a table as shown in Figure 32.



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

Figure 32 Validate sequence listing, verification error table

The location of the HTML file will be displayed alongside the XML verification report (1) & (2) in case the user wishes to copy the files to a different location.

The user must allow an internal script to be run on their machine for the format to display correctly: the ActiveX control on IE browser. This must be done to: 'allow blocked content' for the format to load correctly. Otherwise the sequences will not be displayed in the standard format and will be less readable.

*Please **note** that for validating a Sequence Listing, the ST.26 file should comply with the following requirements:*

- *Must be encoded in UTF-8 and must contain valid characters according to XML 1.0 specification*
- *Must contain a DOCTYPE line as follows:*

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

- *Attribute dtdVersion should be compliant with the DTD version:*
- *The file must comply with DTD file ST26SequenceListing_V1_3.dtd[2].*

4.1.5 Delete Project

To delete a project, the user must begin from the Projects View, shown in Figure 33.

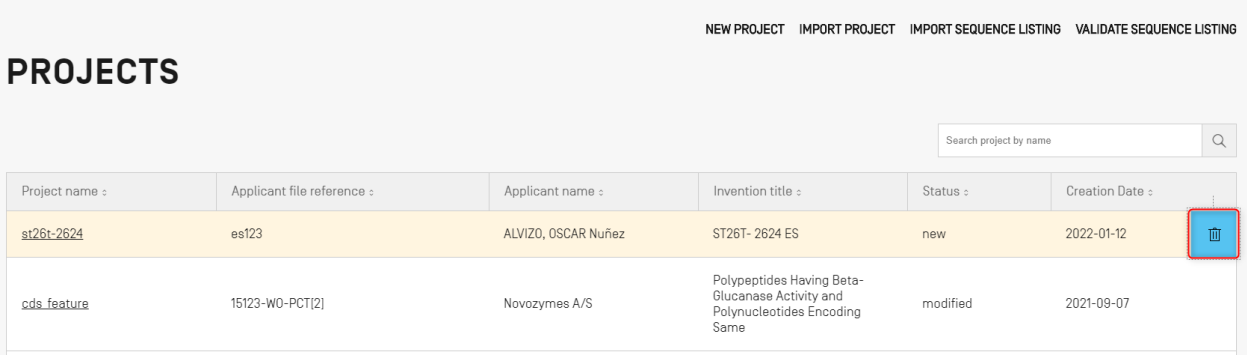


Figure 33 Delete project

Click on the button with the trash can icon on the row within the Projects Home View Table that the user wishes to delete.

[2] This is the current version of the DTD at the time of publication of this document

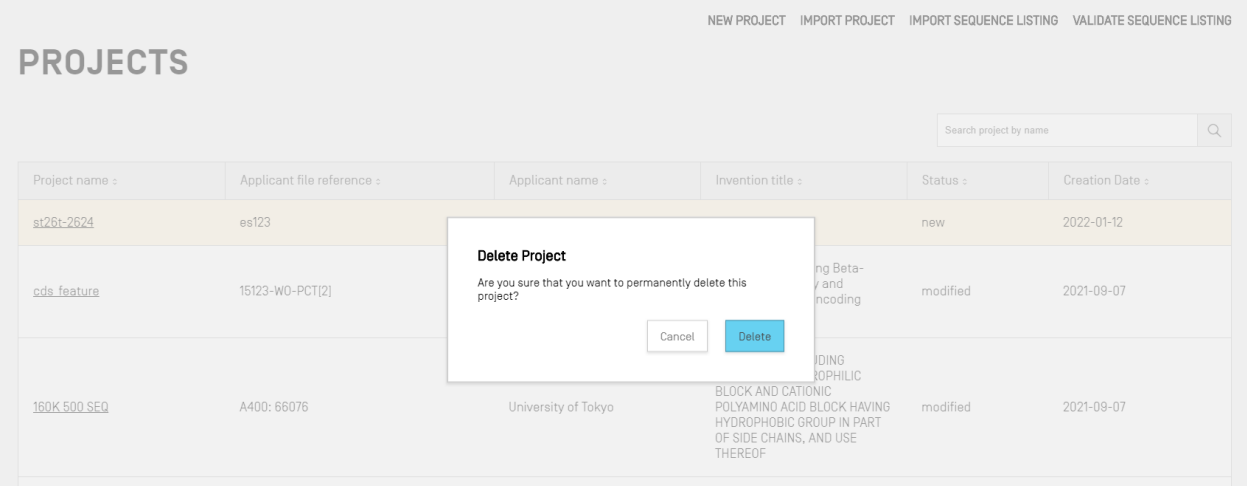


Figure 34 Delete project (confirmation)

In the pop-up click “Delete” to confirm that you want to delete the selected project.

4.2 PERSON & ORGANIZATIONS

4.2.1 Create

To create a new Person or Organization, the user must begin from the Persons & Organizations View.

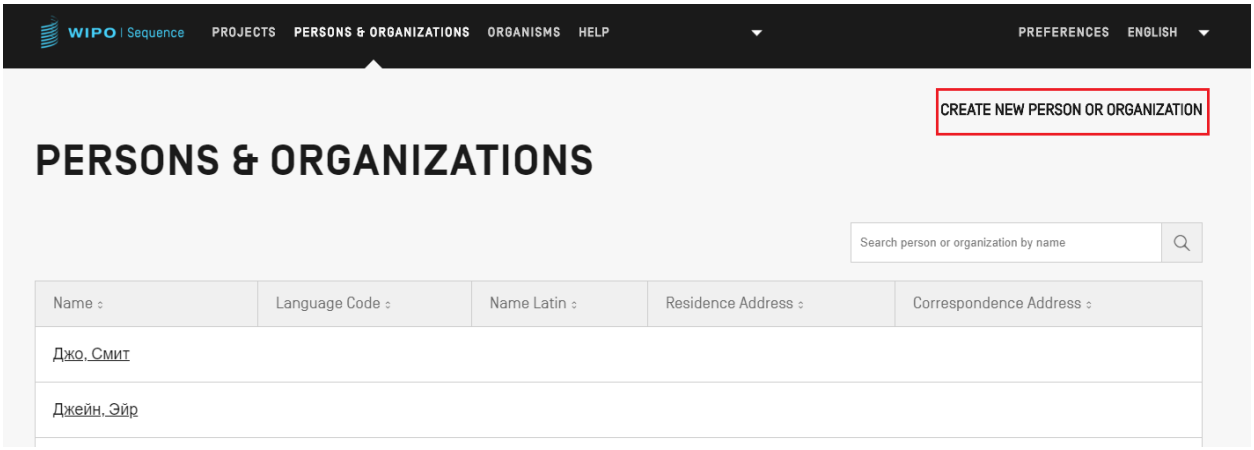


Figure 35 Create new person/organization

First, the user must click on the “CREATE NEW PERSON OR ORGANIZATION” link at the top of the View, as shown in Figure 36.

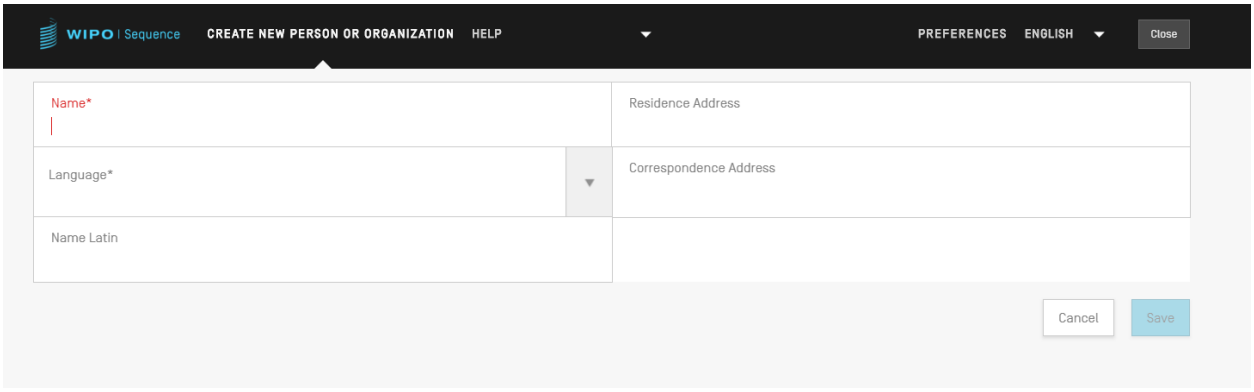


Figure 36 Create new person/organization Panel

In the new View, the user must at least fill in the mandatory fields (indicated with a '*') corresponding to the details of the new person/organization. For the applicant/inventor this is the name (if provided in Latin characters) and the language only.

When the name of the person or organization is not in Latin characters, then the Latin version of the name should be provided in the "Name Latin" field. If this information is not provided then the project will not validate when the ST.26 sequence listing is validated or generated.

4.3 CUSTOM ORGANISMS

To create, edit, import, export or delete Custom Organisms, the user must begin from the Organisms View.

4.3.1 Create Custom Organism

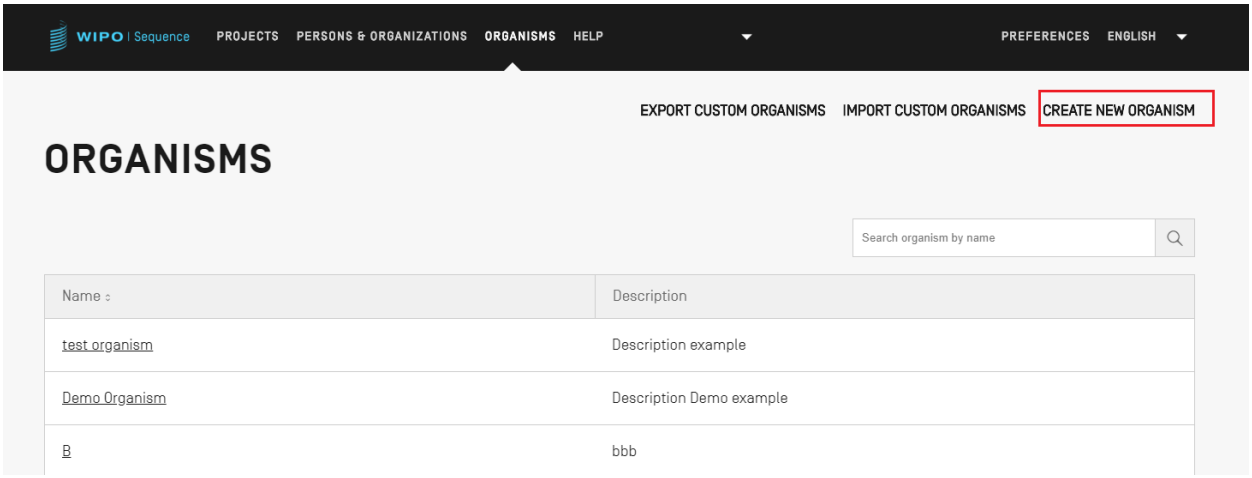


Figure 37 Create new organism

To create a new custom organism, click the “CREATE NEW ORGANISM” link at the top of the view, shown in Figure 37. In the screen that follows (Figure 38), enter the name of the new Organism and click “Save”. If a description of this custom organism is required, this can be optionally added as shown.

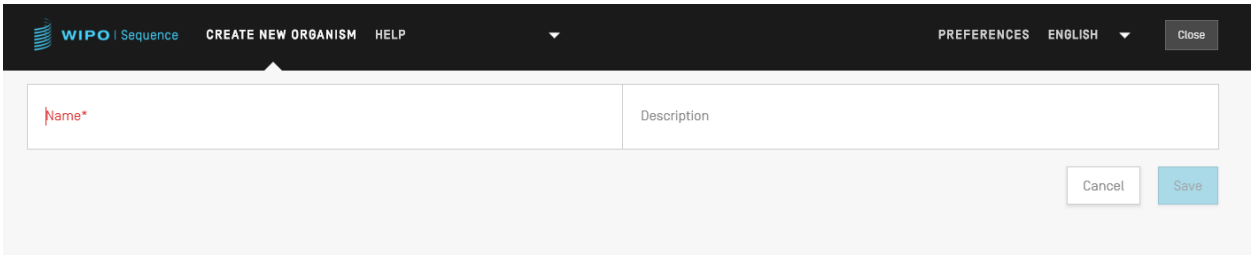


Figure 38 Create new organism (form)

4.3.2 Export Custom Organisms

All the custom organisms and their description that are stored in the tool can be exported and saved to a text file to be modified outside the tool or imported on a later date. To export this list, start by selecting ‘EXPORT CUSTOM ORGANISMS’, as highlighted in Figure 39:

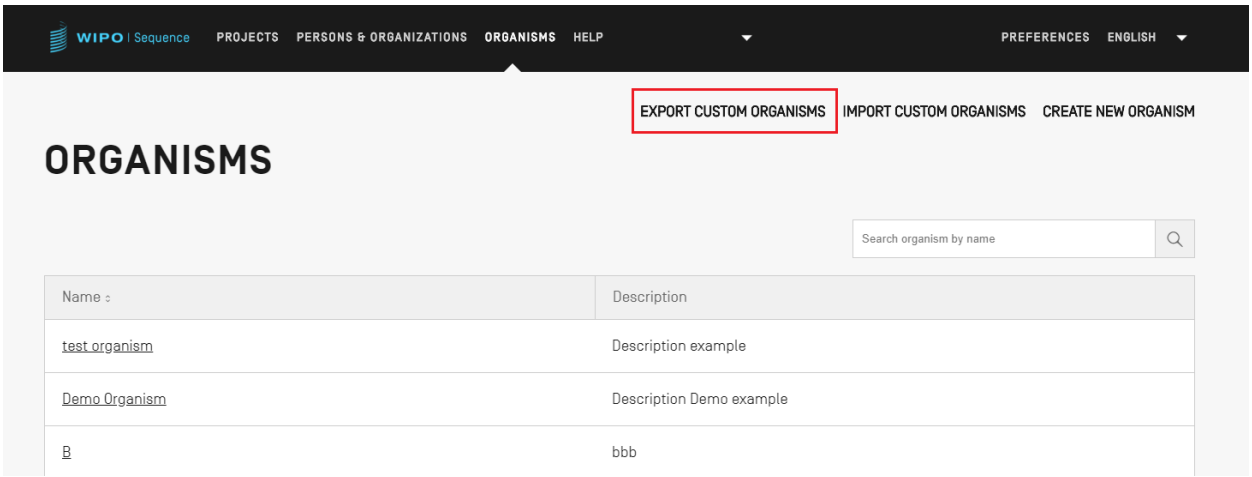


Figure 39 Export custom organisms

Next, a dialog box will open allowing the user to choose the name of the file and the desired file location, see Figure 40.

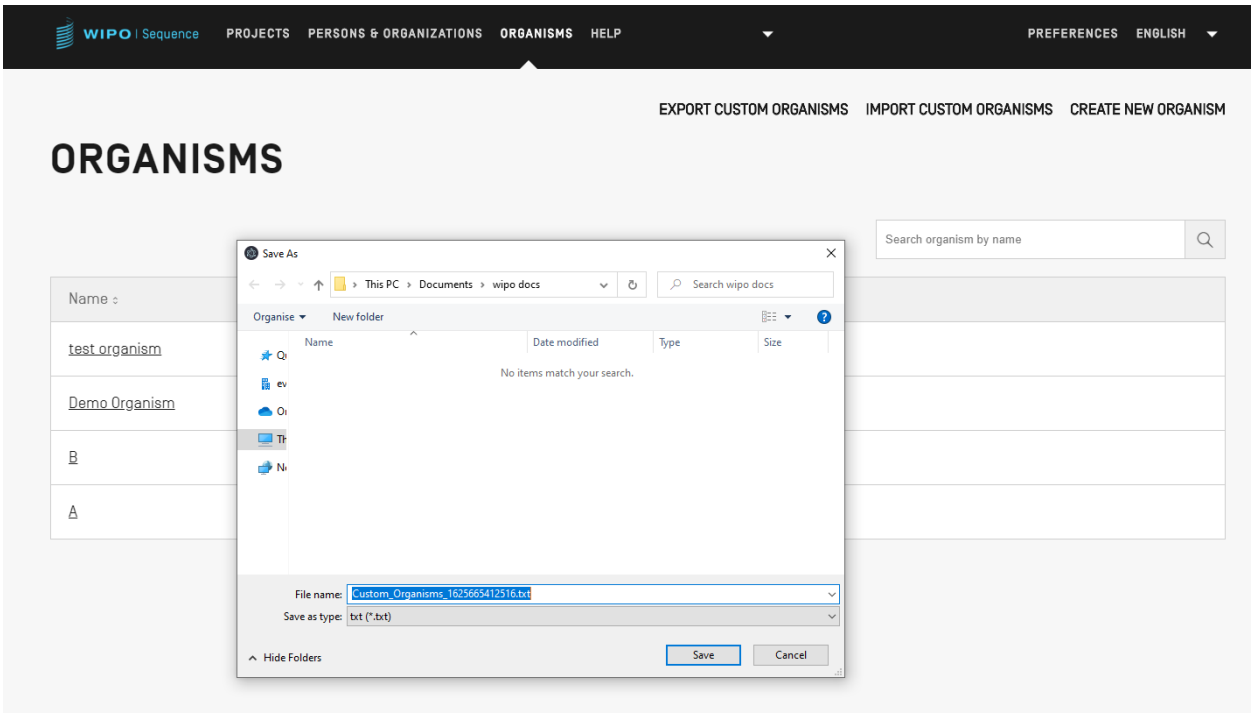


Figure 40 Export custom organisms dialog box

The file that is exported is txt file including both the name and the description of the organism which could be Edited and imported into the tool (see Figure 41).

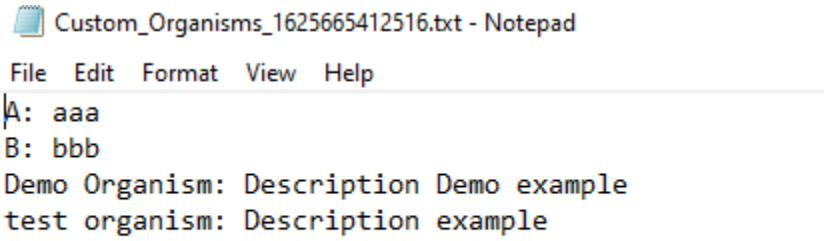


Figure 41 Example of Custom organism file

4.3.3 Import Custom Organisms

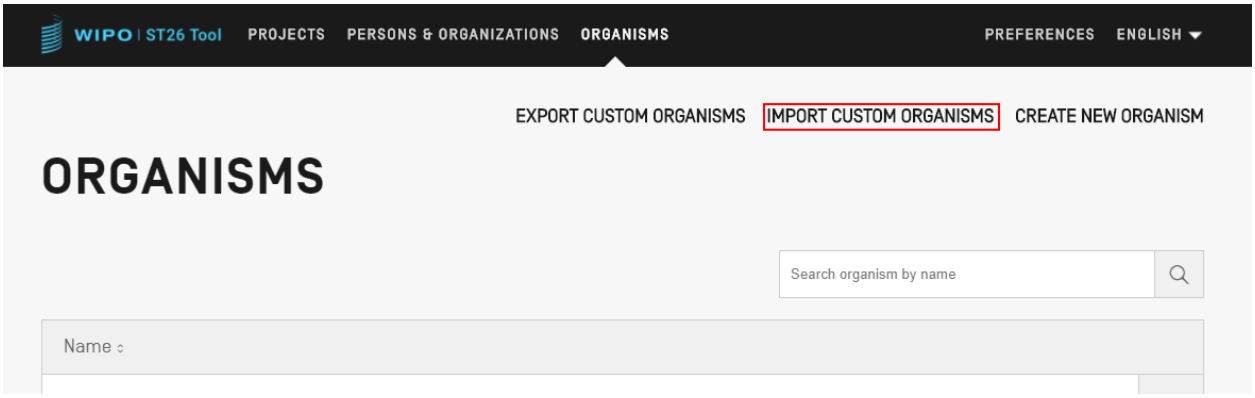


Figure 42 Import custom organisms

Firstly, in order to import a list of custom organisms, the user must click on the “IMPORT CUSTOM ORGANISMS” link at the top of the View, as shown in Figure 42. This will open an Overlay below the Organisms Table, as shown in Figure 43.

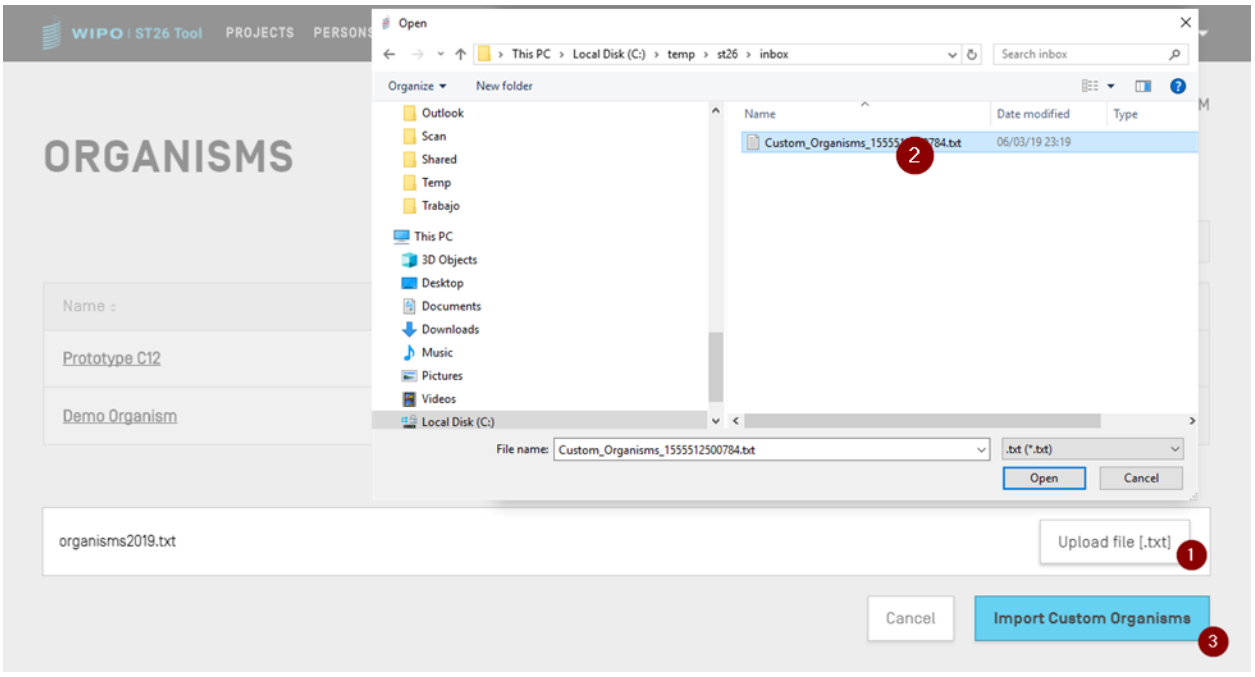


Figure 43 Import custom organisms (dialog box)

1) Click on the “Upload file [.txt]” button (1)

2) Select the file with the custom organism names from within the dialog box (2)

3) Finally, click on the blue “Import Custom Organisms” button (3)

Note:

The file to be imported will be a text file (.txt) with a list of custom organism names in plain text (UTF-8), each item on a new line.*

4.4 SYSTEM PREFERENCES

The System Preferences View allows the modification of several configuration parameters of WIPO Sequence. These parameters will apply to every project created or edited by the tool (see Figure 44).

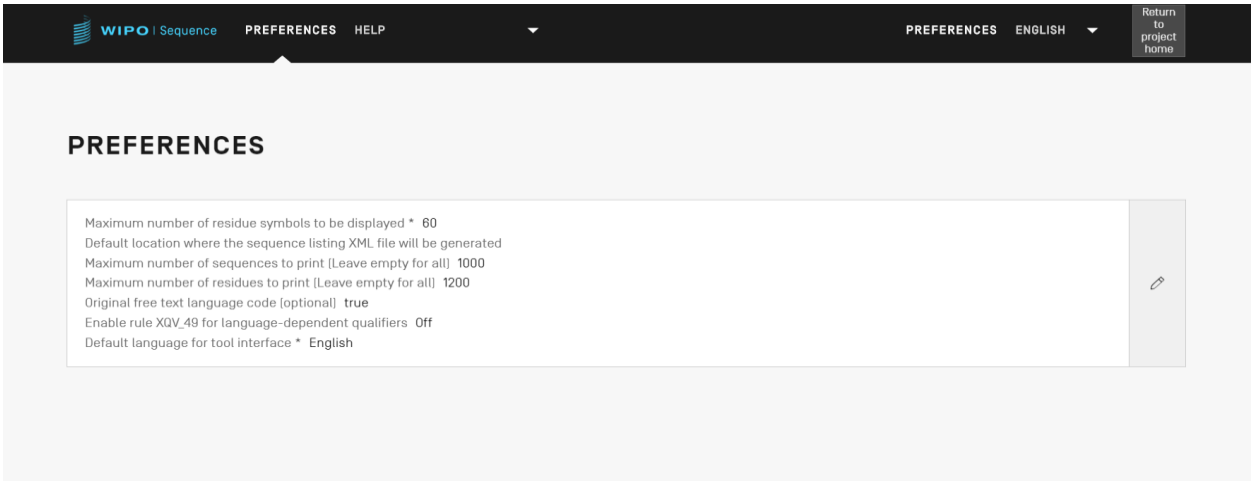


Figure 44 Summary of System Preferences

In order to modify the system preferences, the user should click on the pencil icon shown above to open the Edit Panel shown in Figure 45:

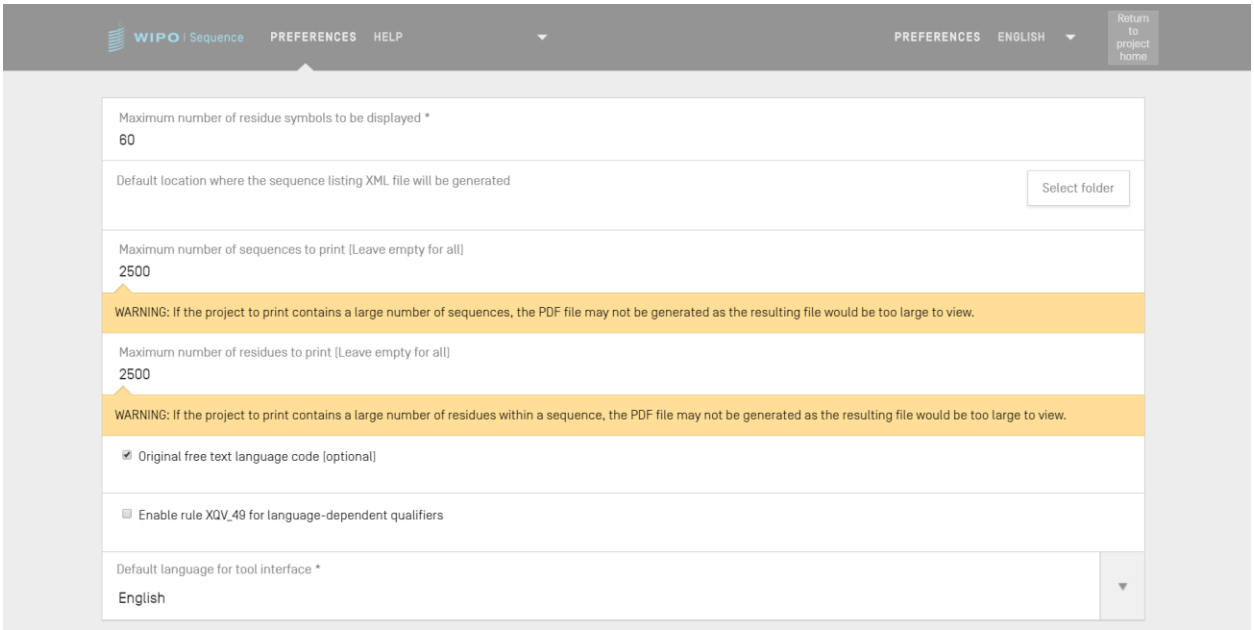


Figure 45 Configuring system preferences

The list of configuration items that can be modified from this View (in order) are:

- Maximum number of residue symbols to be displayed: This parameter sets the number of residues that will be displayed per row when displaying a sequence. The default is 60 residues.
- Default location where the ST.26 sequence listing file (.xml) will be generated. There is no need to provide this location.
- Maximum number of sequences to print (leave empty for all): the default is 1000 sequences.
- Maximum number of residues to print (leave empty for all): the default is 1200 residues.
- Original Free Text language code: If this checkbox is checked, then a warning will be thrown during validation is the original free text language code is not provided. By default, this is unchecked.
- Enable XQV_49 : If this checkbox is checked, then a warning will be thrown if there is no English value for a language dependent free text qualifier provided. By default this is unchecked.

- Default interface language: This is the language in which the interface will appear when WIPO Sequence is launched. By default this is English

Note:

The the third and forth items are relevant when Printing the project as a PDF. Users should notes that for very large sequence listings, the resulting PDF can have several thousand pages and be impossible to display.

5 PROJECT PAGE

The project page is composed of six Views that can be used to navigate between different portions of the workflow, as shown in Figure 46:

1. Project Detail View (1) (indicated by the name of the project, shown here as 'Project'): main view which contains all of the project data,
2. Verification Report View (2): where the verification report can be accessed,
3. Language Dependent Qualifiers View (3): where the language dependent free text qualifiers can be accessed and exported/imported,
4. Import Report View (4): where the import report can be accessed,
5. Display Sequence Listing View (5): where human readable formats of the generated ST.26 sequence listing can be accessed and
6. Help Menu: includes references to the user manual and WIPO ST.26
7. Preferences View (6).

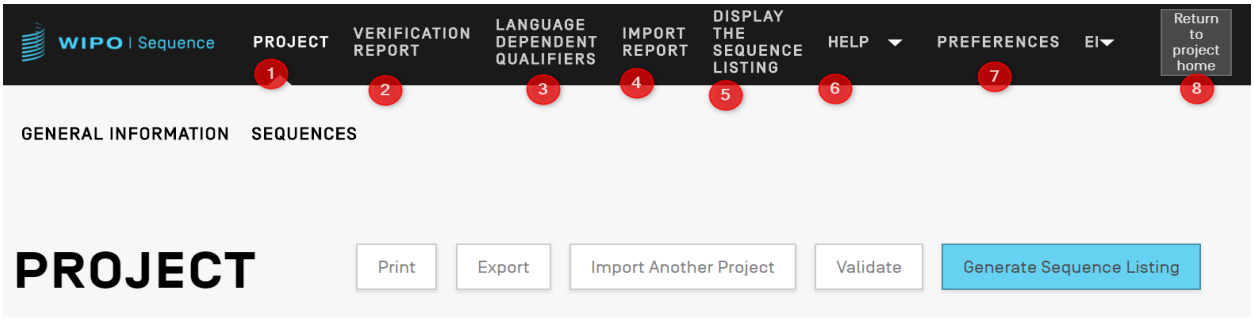


Figure 46 Project page header

To return to the Projects View (main home page), the user can click on the “return to project home” button (8) at the right end of the header.

5.1 PROJECT DETAIL

5.1.1 Basic Information

A Table containing the basic information about the project can be found at the top of the Project Detail View, shown in Figure 47.

This section contains:

- Name of Project
- Date and time of creation of the project
- Project status (possible values: ‘new’/’modified’/’generated’/’invalid’/’valid’/’warnings’) - note this is not an editable field
- Name of the imported file (in the case that the project was imported)
- Project description - optional
- Number of Sequences (labelled: ‘Sequences’)
- Original free text language code for free text qualifiers
- Non English free text language code
- A checkbox for invoking the automatic addition of a translation qualifier when a CDS feature is created (a project-level function)

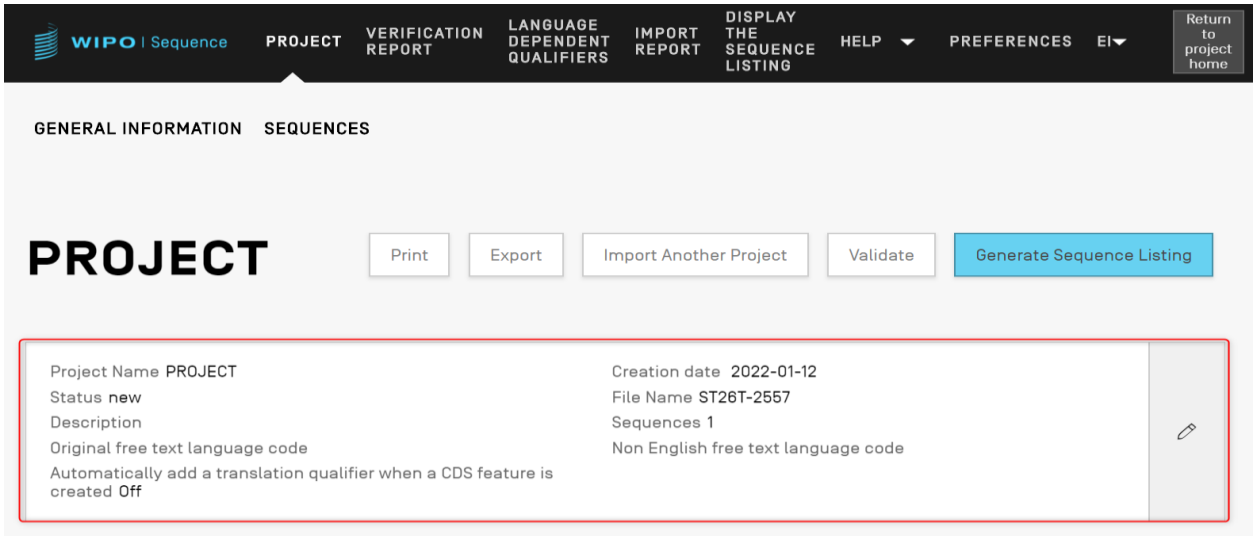


Figure 47 Basic Information Section

5.1.1.1 Print a project

To print a project, the user must enter the Project Detail View of the desired project and click on the “Print” button at the top of the view, see Figure 48.

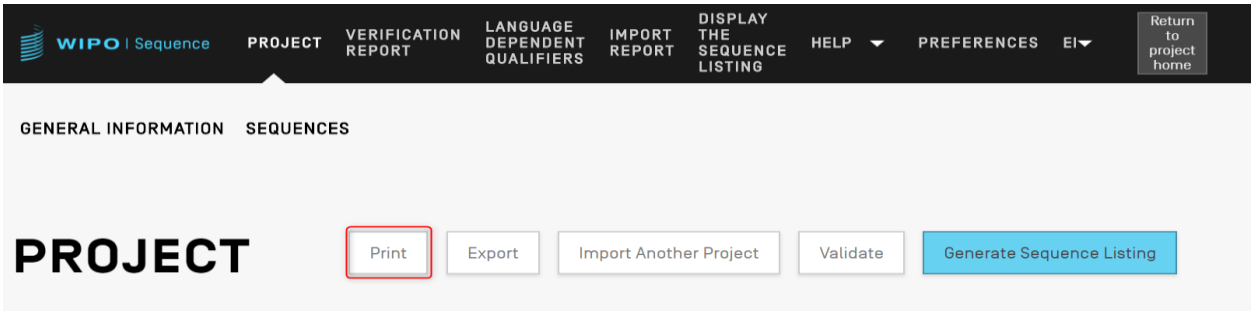


Figure 48 Print project

Next, the user will be shown two checkboxes to clarify what information the user wants to print from the project: General Information and/or Sequence Information (see Figure 49).

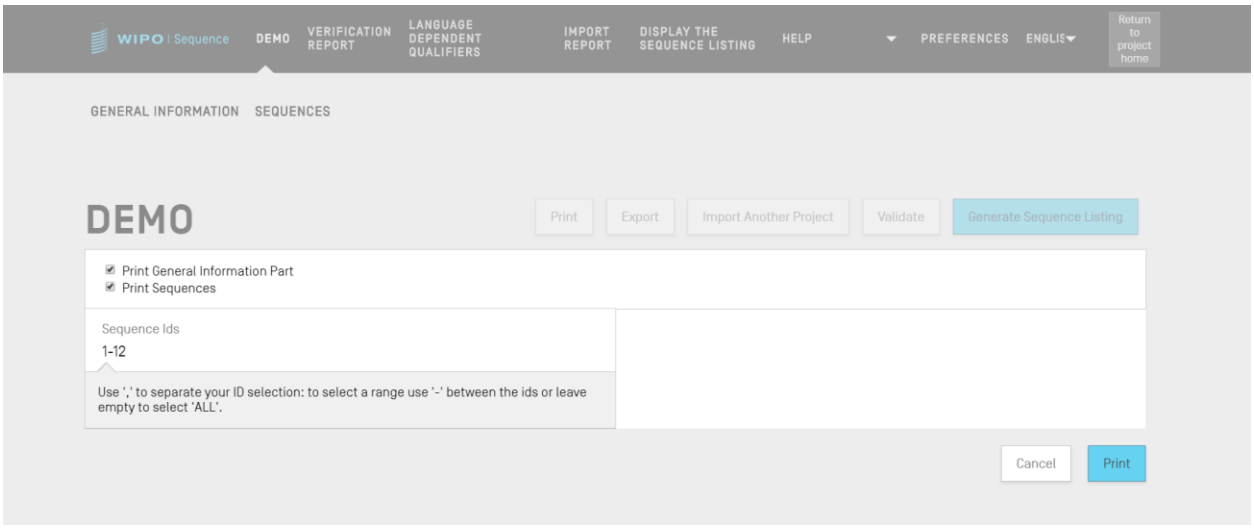


Figure 49 Print project, Section selection

If “Print Sequences” is selected, the user will have the choice to specify which sequences are to be printed by specifying the range of ID numbers within the “Sequence IDs” field, or simply print all if this field is left blank.

By default, the total number of sequences of the project will be displayed as a range.

Once the blue “Print” button has been clicked, if the PDF file was generated correctly, the tool will open the file in a PDF reader for the user’s review.

For instructions on how to download the PDF file, see Section **3.2.**

Using any internet browser, the user may also save the displayed HTML format sequence listing as a PDF file.

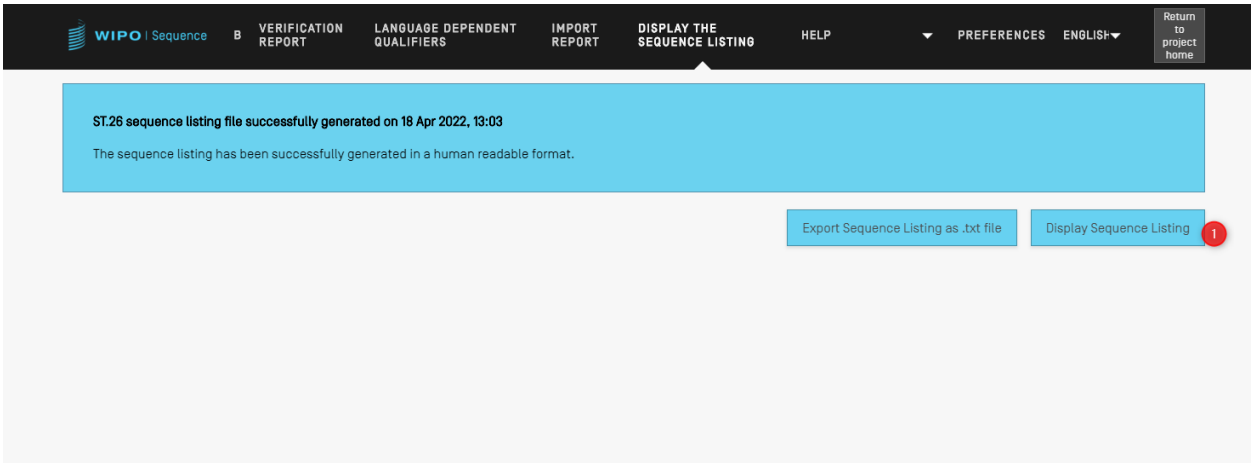
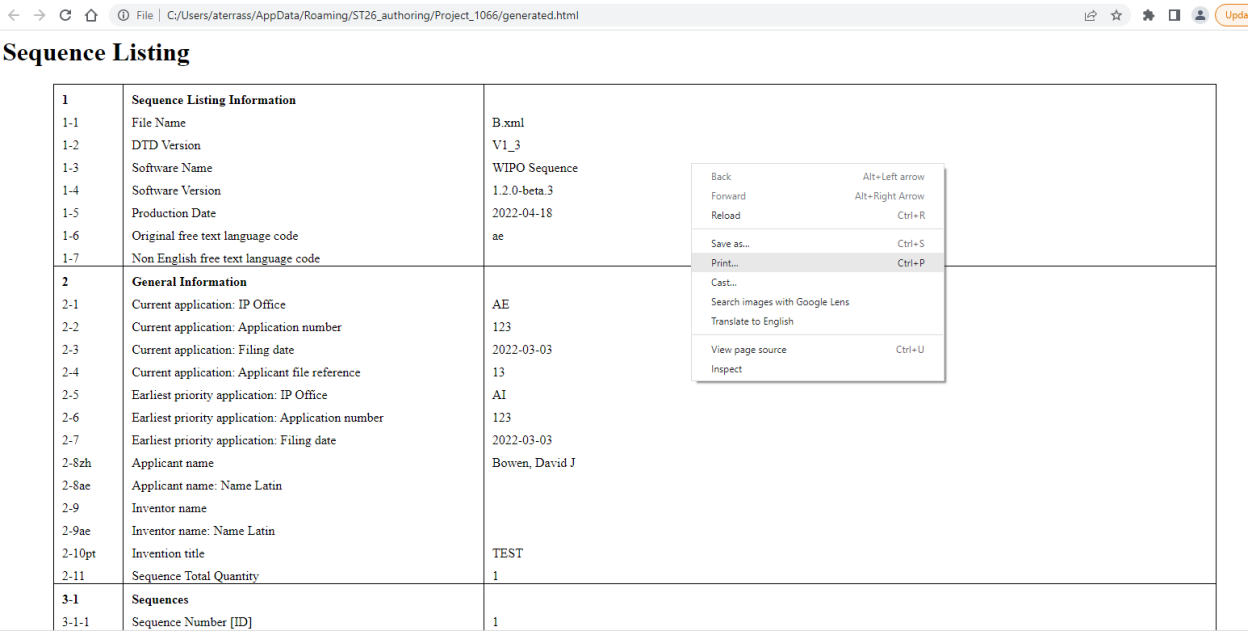


Figure 50 Display Sequence listing



1	**Sequence Listing Information**	
1-1	File Name	B.xml
1-2	DTD Version	V1_3
1-3	Software Name	WIPO Sequence
1-4	Software Version	1.2.0-beta.3
1-5	Production Date	2022-04-18
1-6	Original free text language code	ae
1-7	Non English free text language code	
2	**General Information**	
2-1	Current application: IP Office	AE
2-2	Current application: Application number	123
2-3	Current application: Filing date	2022-03-03
2-4	Current application: Applicant file reference	13
2-5	Earliest priority application: IP Office	AI
2-6	Earliest priority application: Application number	123
2-7	Earliest priority application: Filing date	2022-03-03
2-8zh	Applicant name	Bowen, David J
2-8ae	Applicant name: Name Latin	
2-9	Inventor name	
2-9ae	Inventor name: Name Latin	
2-10pt	Invention title	TEST
2-11	Sequence Total Quantity	1
3-1	**Sequences**	
3-1-1	Sequence Number [ID]	1

Figure 51 Sequence listing as PDF

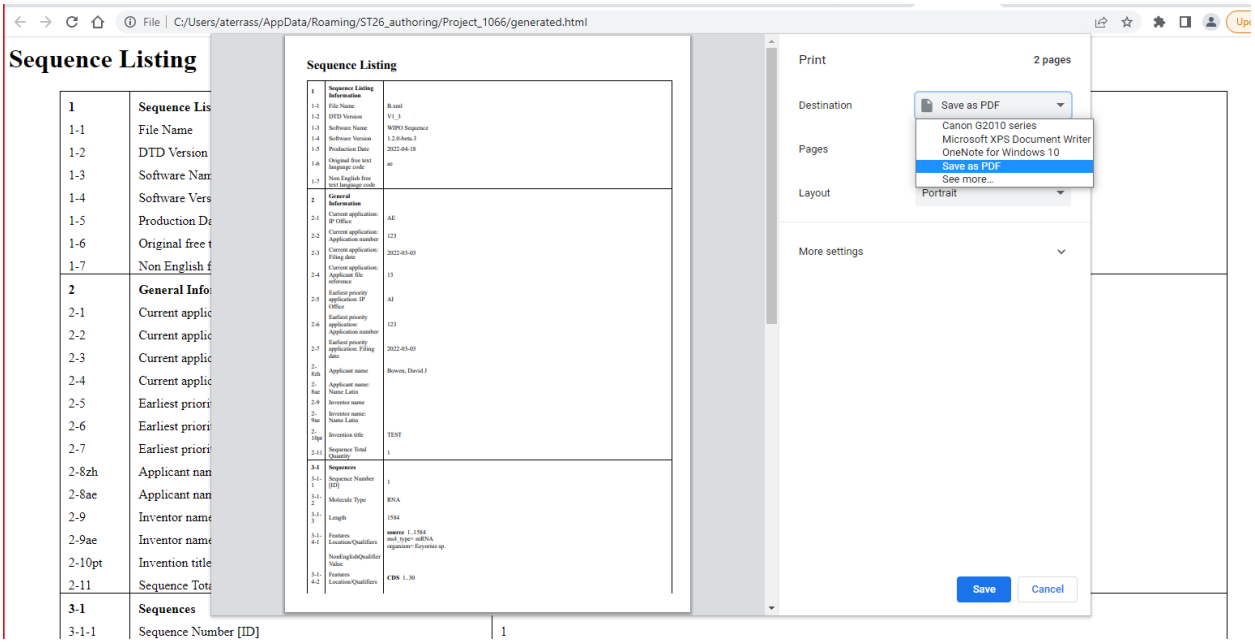


Figure 52 Save Sequence listing as PDF

5.1.1.2 *Import Information From Another Project*

The user can copy information from other projects stored in the tool, into the currently open project. This imported information can be either for the “General Information” Section, “Sequences” Section, or both.

Note: Imported General Information will replace the currently existing General Information in the project, while imported Sequences will be appended to the current list of sequences within the project.

To begin, the user must click on the “Import Another Project” button at the top of the Project Detail View, see Figure 53.

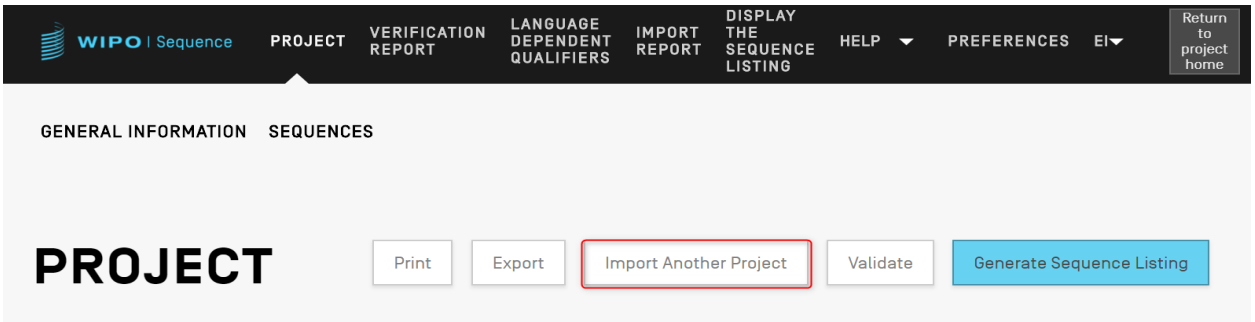


Figure 53 Import Another Project

The tool will open an Overlay, as shown in Figure 54. The user must first select the project from which they wish to import information.

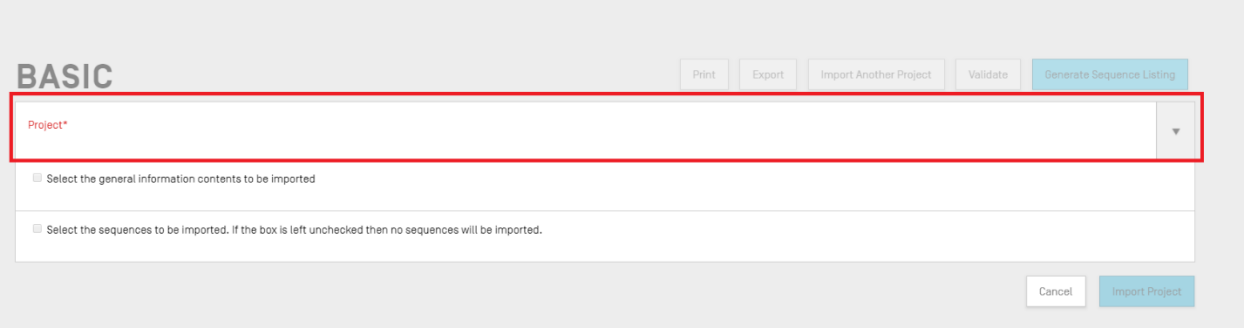


Figure 54 Import Another Project, select target

Next, as shown in Figure 55, the user can select whether they wish to include parts of the details provided in the General Information (1) Section of the project and also if they wish to import sequences (2) by providing range of sequence ID numbers (3) to specify which of the sequences are to be imported into the project.

By default, the total number of sequences of the project will be displayed as a range.

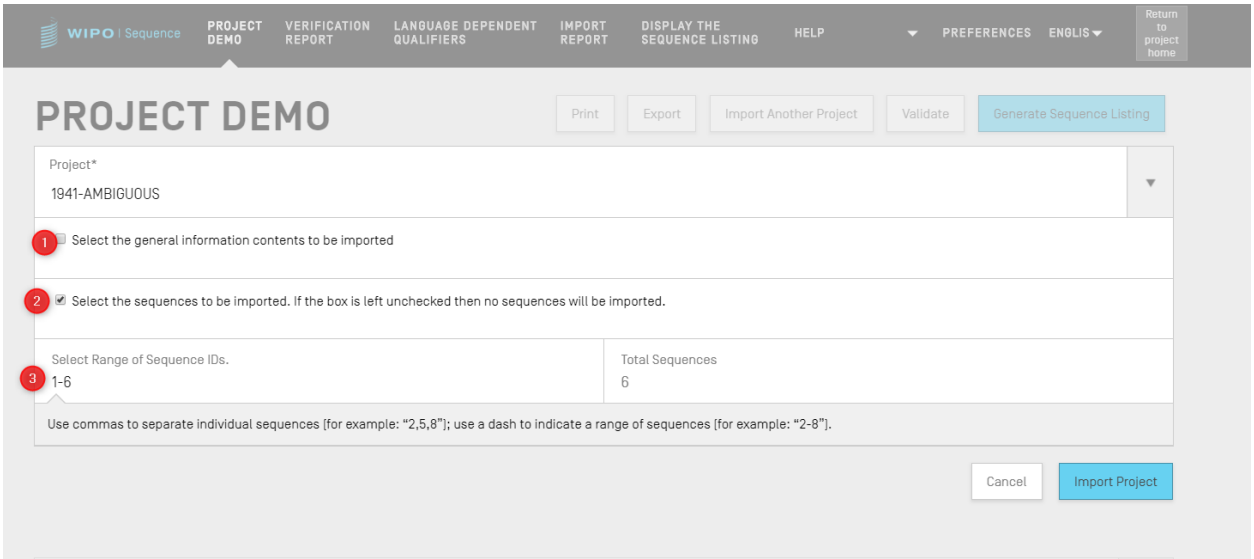


Figure 55 Import Another Project, Panel

If the General Information checkbox is checked then a Table will appear displaying all the General Information Section of both projects: the currently selected (origin) project, and the target project (destination).

Note:

if the user does not check the 'Select the general information contents to be imported' box, then no general information from sequence listing will be imported into the project.

The user must then select which of the General Information elements are to be replaced by the corresponding target project's General Information, as shown in Figure 56.

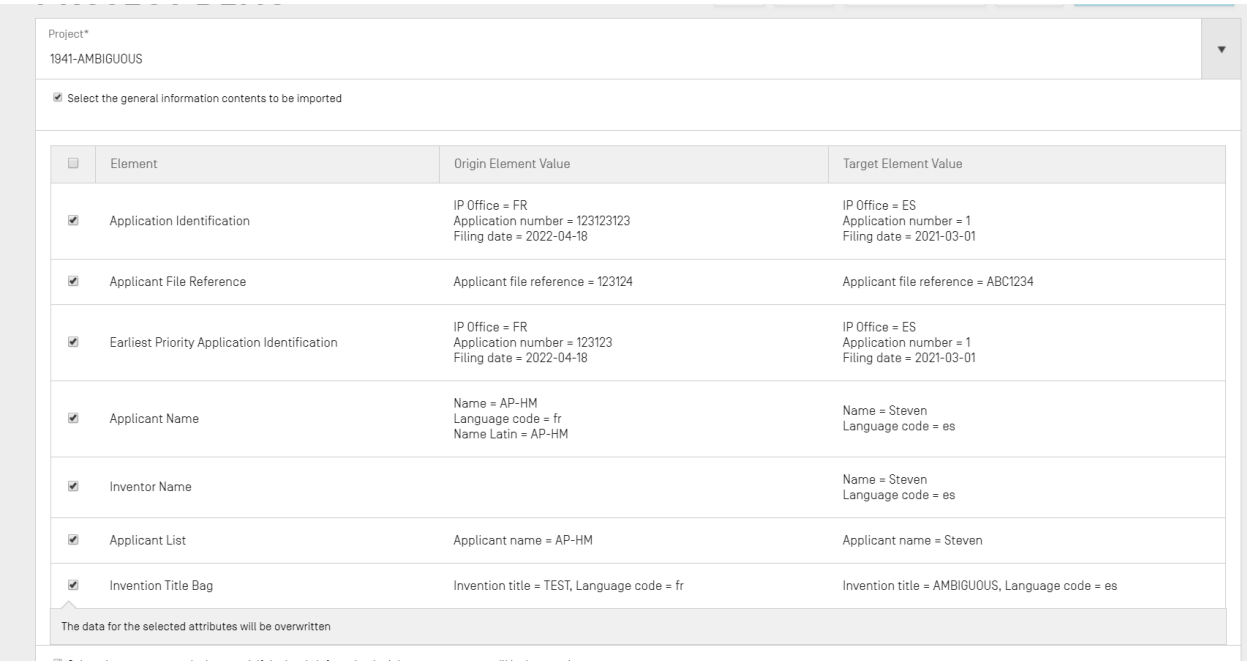


Figure 56 Import Another Project, select general information elements

Finally, when the user has decided on which General Information elements and sequences are to be imported into the project, then the user must click on the blue "Import Project" button, shown at the bottom of Figure 56.

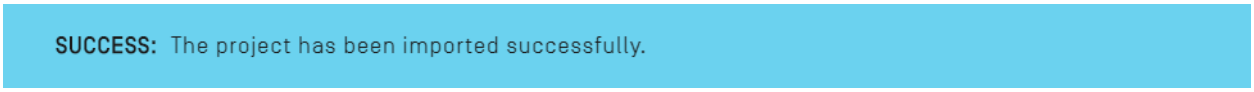


Figure 57 Import Another Project, success banner

As shown in Figure 57 Import Another Project, success banner, a blue banner appears if the elements have been imported correctly.

5.1.1.3 Validate Project

Before generating the sequence listing as an ST.26-compliant XML file, a project will pass through a validation check, beforehand. This step is always conducted prior to generating the sequence listing but can also be performed on its own.

To validate a project, the user must click on the “Validate” button at the top of the Project Detail View, shown in Figure 58.

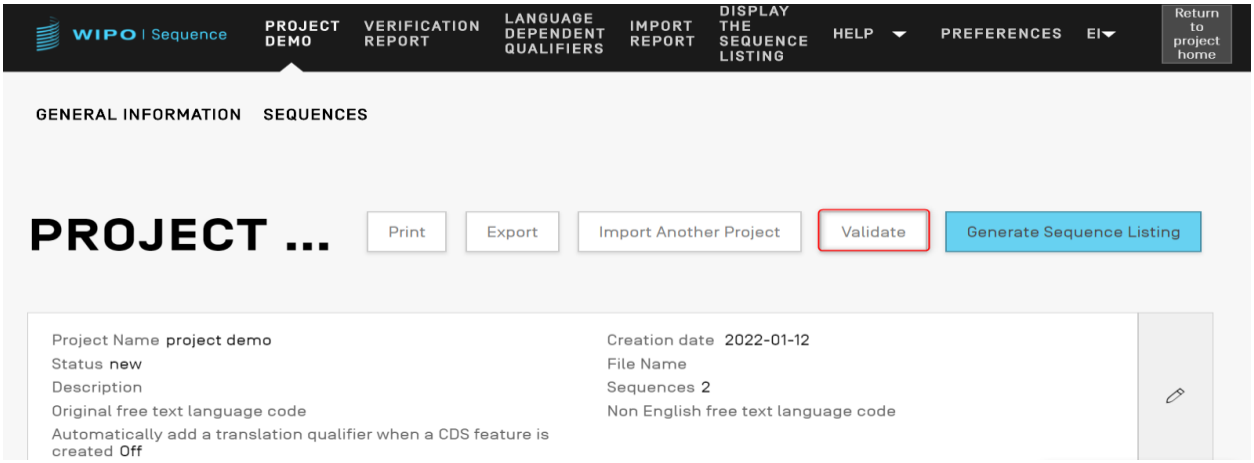


Figure 58 Project validation

Once the validation has finished, the user will be brought to the “Verification Report” View, displaying any the verification errors/warnings that may be generated. Figure 59 shows the screen that will be displayed in the case of successful validation.

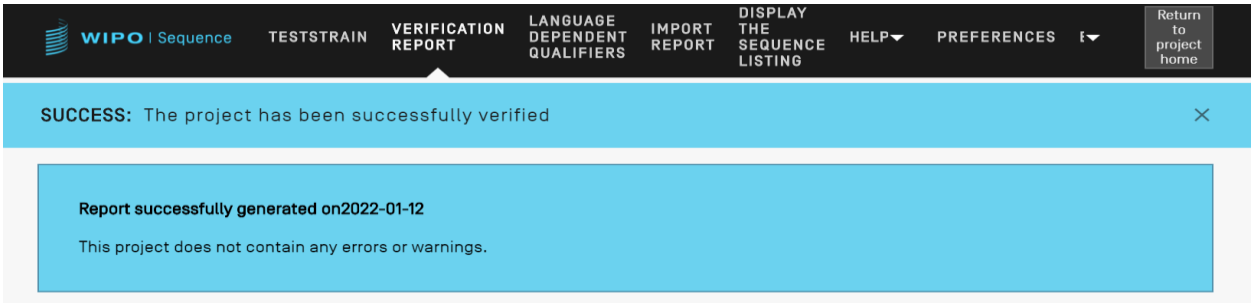


Figure 59 Project validate, no errors

If the validation process finds any errors or warnings, a Verification Report will be generated with a Table detailing the detected verification rules and guidelines that have been broken. An example report is shown in Figure 60. Each row is also identifying whether this is an error, which must be addressed, or a warning, must can be ignored by the user.

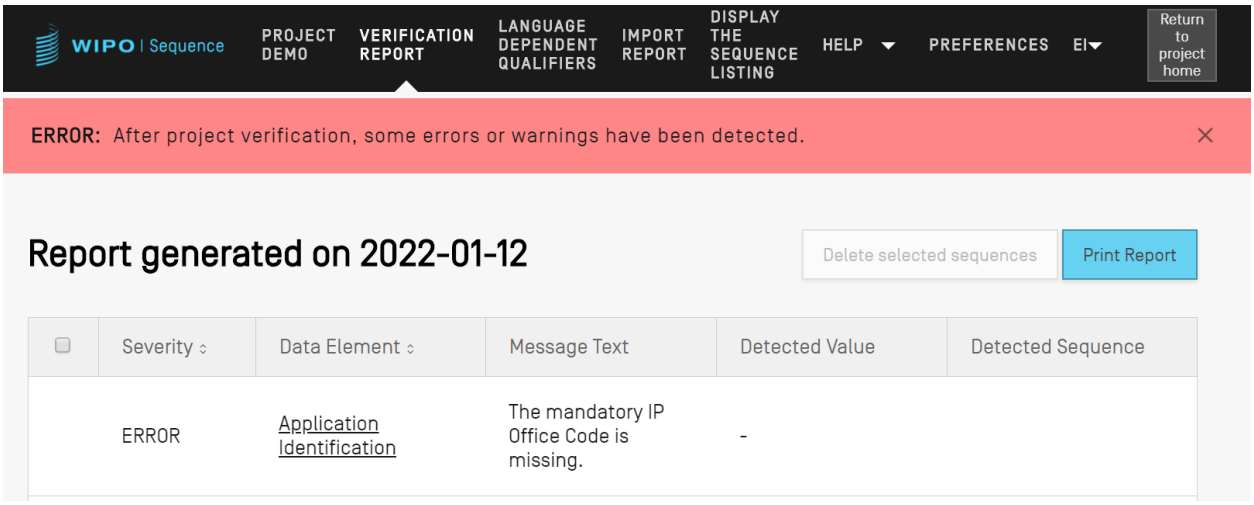


Figure 60 Project validate, errors/warnings

5.1.1.4 Export Project

A project can be exported to a .zip file for the user to back-up project data or alternatively import with another desktop computer with WIPO Sequence installed, as shown in Figure 61.

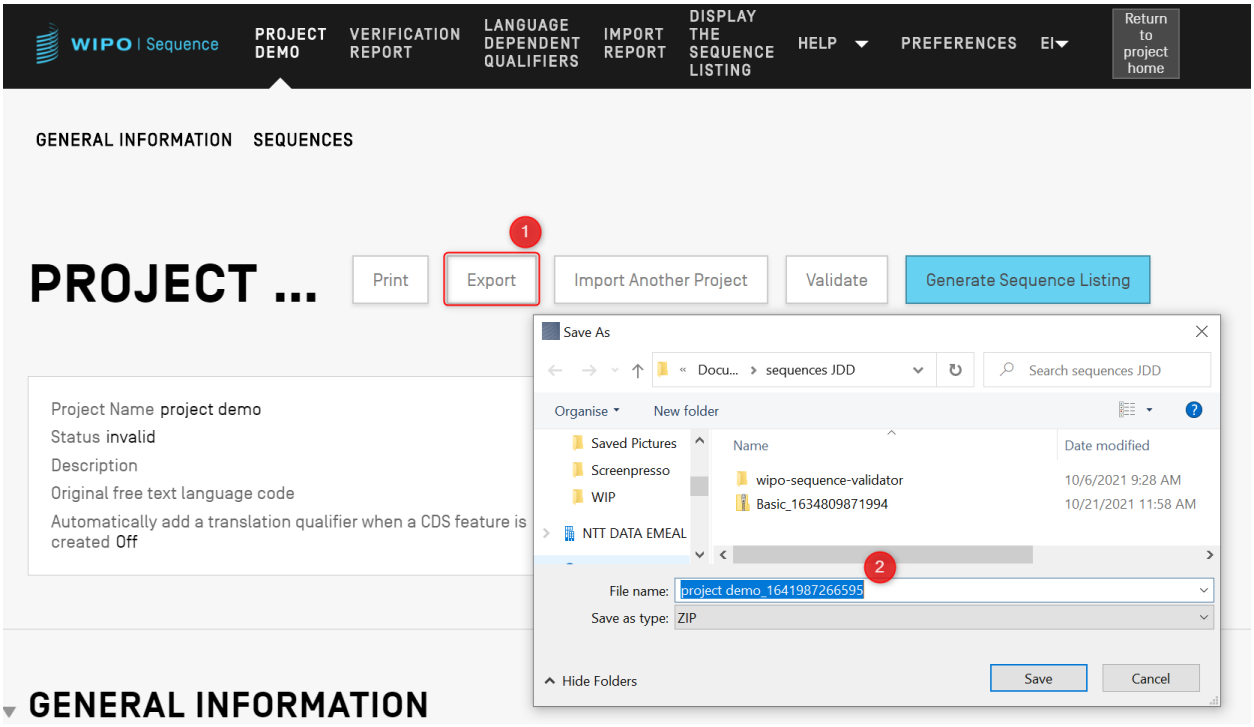


Figure 61 Export Project

1) Click on the “Export” button at the top of the Project Detail view.

2) In the dialog box that appears, select the file name and desired location to save the project.

If the project is successfully exported, a blue banner with the saved name and location will appear at the top of the screen as shown in Figure 62.

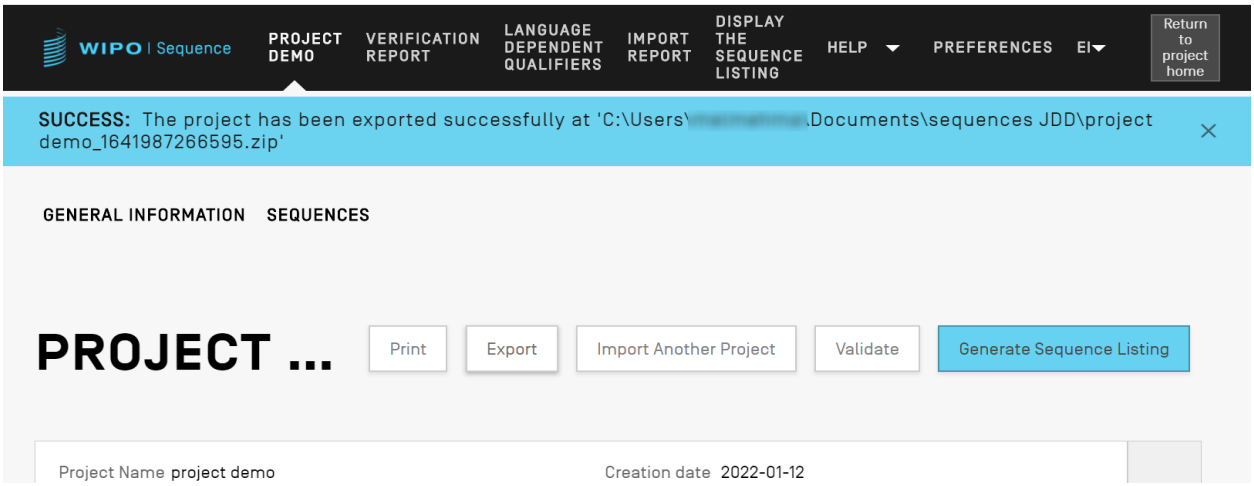


Figure 62 Export Project, success

5.1.1.5 Generate Sequence Listing

The final action that can be performed on a project, and perhaps the most important, is to generate the sequence listing. To generate the sequence listing, the user must click on the blue “Generate Sequence Listing” button, at the top of the Project Detail view, highlighted in Figure 63. This will automatically trigger the validation process to be run on the project first (see **5.1.1.3**).

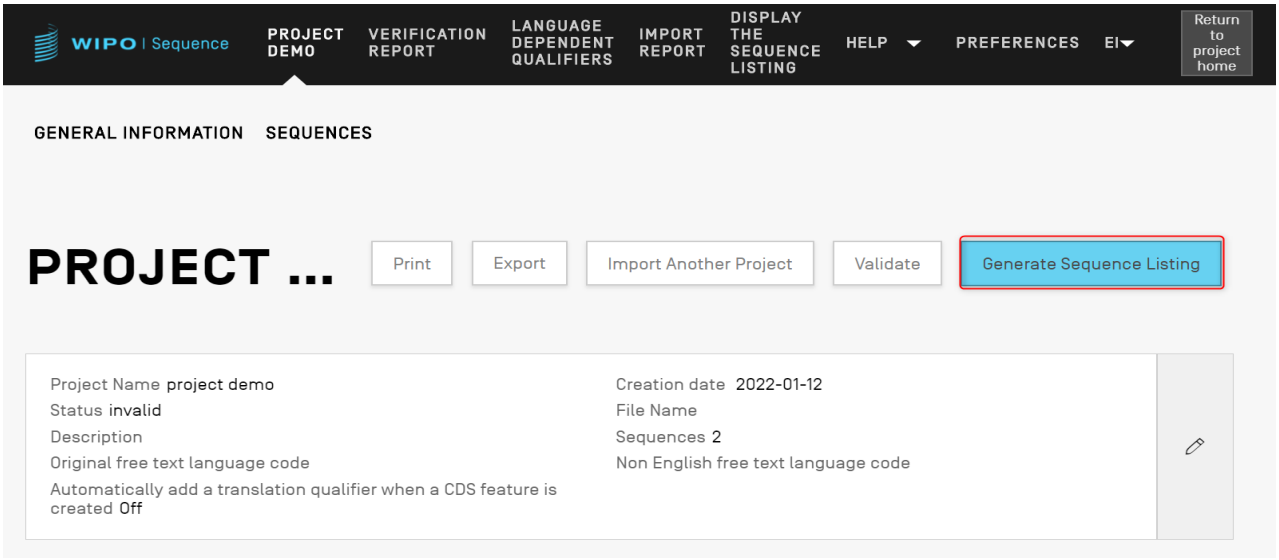


Figure 63 Generate Sequence Listing

If the project passes the validation process, a dialog box will open for the user to select where to save the generated ST.26 compliant sequence listing (.xml), as highlighted in Figure 64.

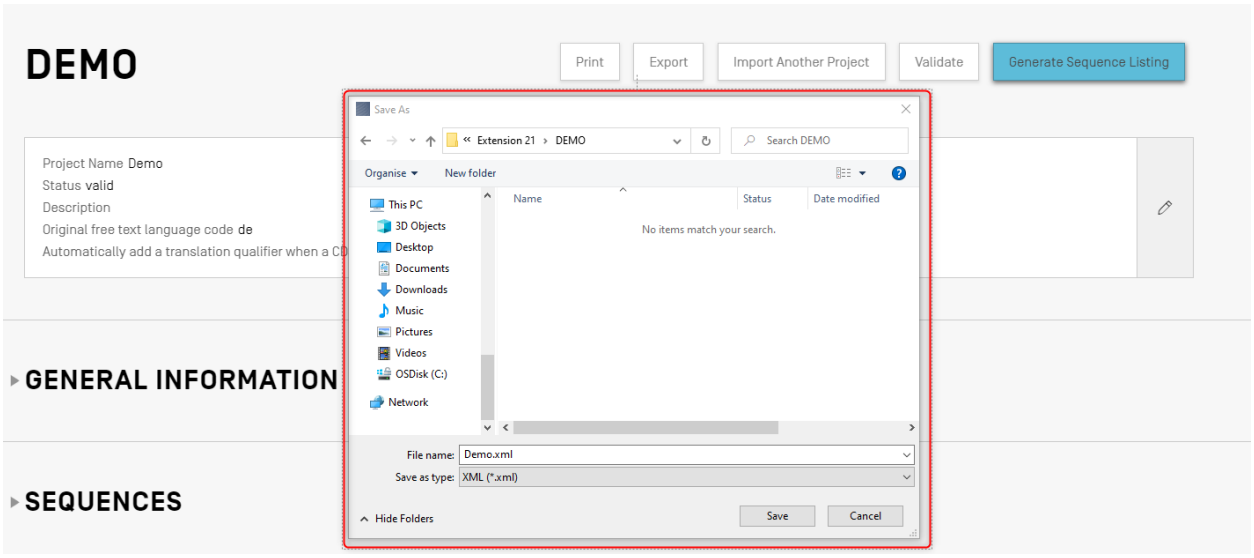


Figure 64 Generate Sequence Listing Dialog box

Note:

There is a known issue when using the Linux distribution: an extra '\' appears in suggesting file name by default. In order to resolve this, please manually remove the extra '\' before saving.

If the project fails validation, then the Verification Report View will instead be presented along with a red banner as shown in Figure 65.

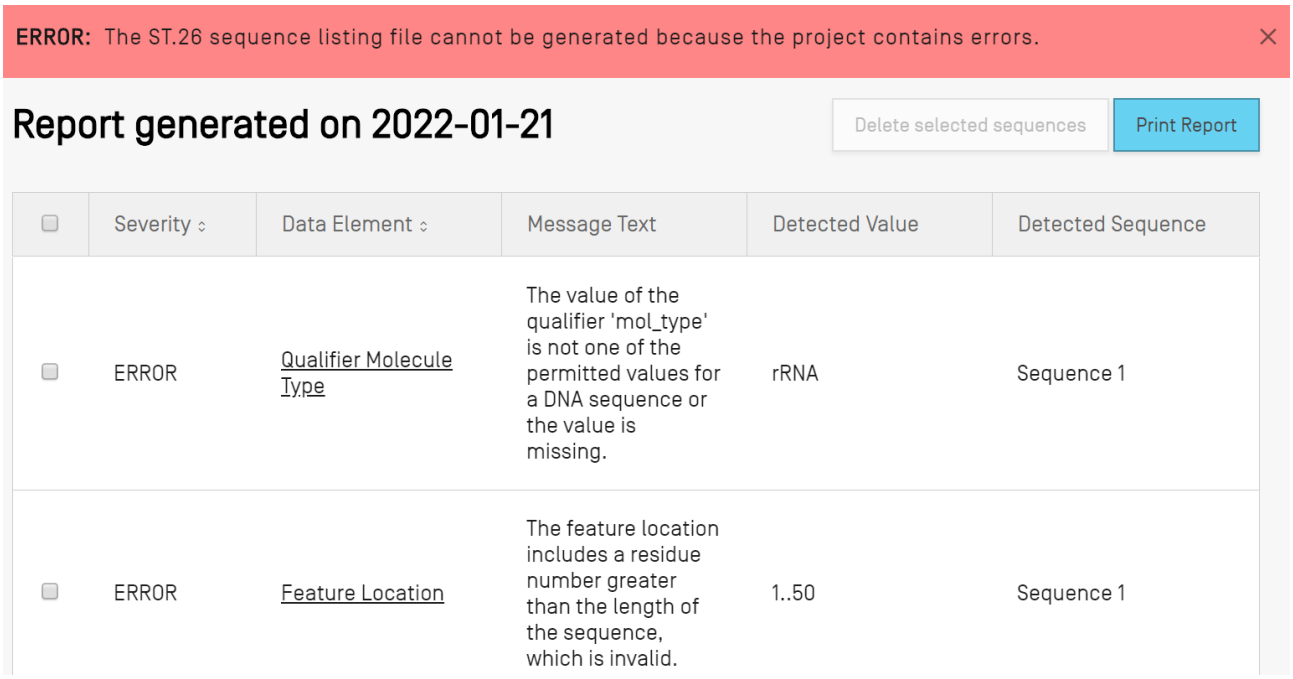


Figure 65 Generate Sequence Listing, Failed

5.1.2 General Information

This Section allows the user to enter information related to the patent application itself, which is used to associate the generated sequence listing with this application.

5.1.2.1 Application Identification

The “Application Identification” subsection of the General Information Section is related to the patent application status and information of the selected project, see Figure 66.

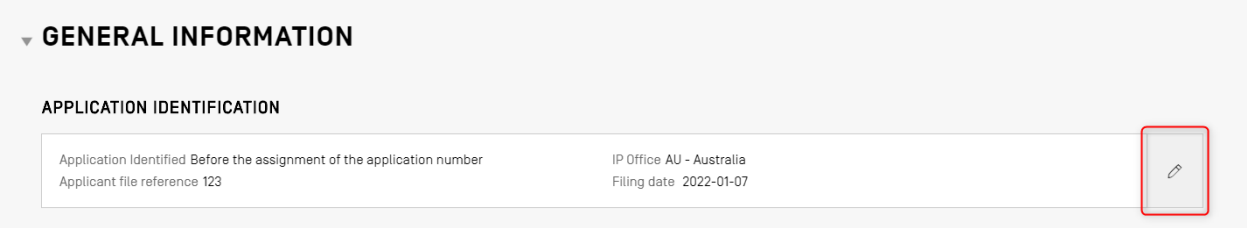


Figure 66 Application Identification

To edit information within the Application Identification subsection, click on the pencil icon highlighted, to the right of the subsection. Then the user must provide information based on the following steps shown in Figure 67:

1) If the application already has an assigned application number, the user must select the code of the Intellectual Property Office (IP Office) at which the application was filed. This is the WIPO ST.3 code.

2) The user must select whether or not they have already been notified of the application number or else just provided within an application file, by selecting the appropriate radio button.

3) In the case of not having the application number, the user MUST provide the applicant file reference in this field.

4) If an application number has already been assigned, the user should enter the application number provided for the patent.

5) Select the filing date of the application with the Date Picker if a date has been assigned.

6) Click the blue “Save” button.

Note:

Regardless of what is entered, a warning will always appear in the verification report indicating that "The application identification number is absent. The application number is mandatory if the application number has been assigned."

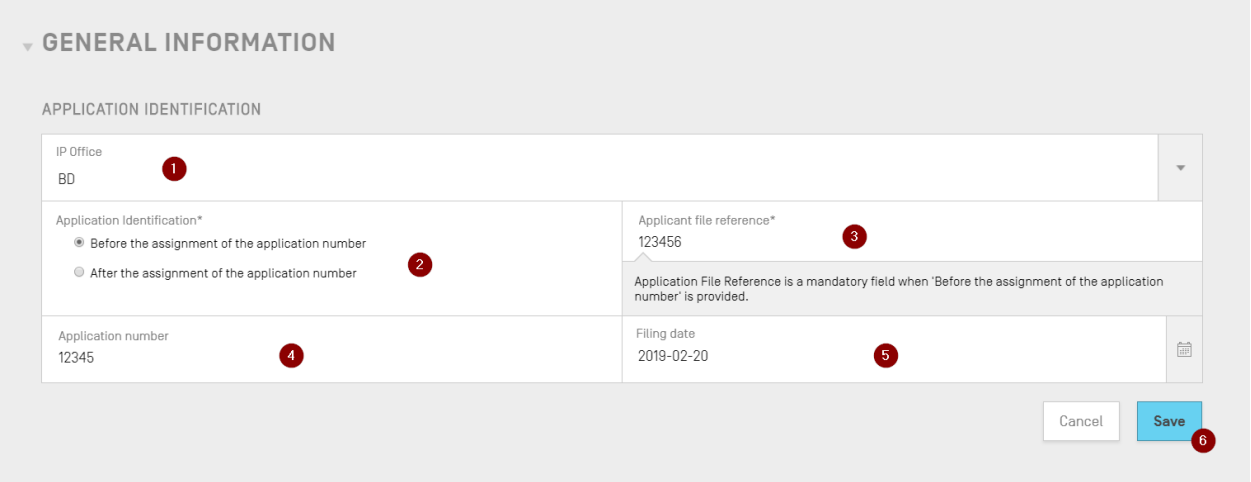


Figure 67 Application Identifiifcation Edit Panel

5.1.2.2 Priority Identification

To add a priority application to the project, the user must click on the “Add Earliest Priority Identification” button in the General Information Section of the Project Detail View, shown in Figure 68.

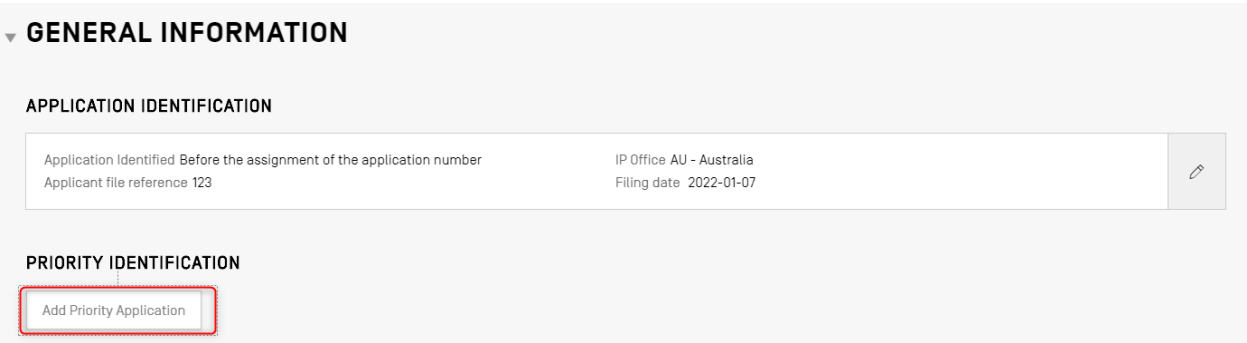


Figure 68 Earliest Priority Application

To set the currently selected priority application as the earliest, the user must select “Yes” in the “Selected Earliest Priority Application” dropdown. This will set or modify this as the priority application established as the earliest priority application when the sequence listing is generated.

To finish, click on the blue “Add Earliest Priority Application” button in the Overlay, shown in Figure 69.

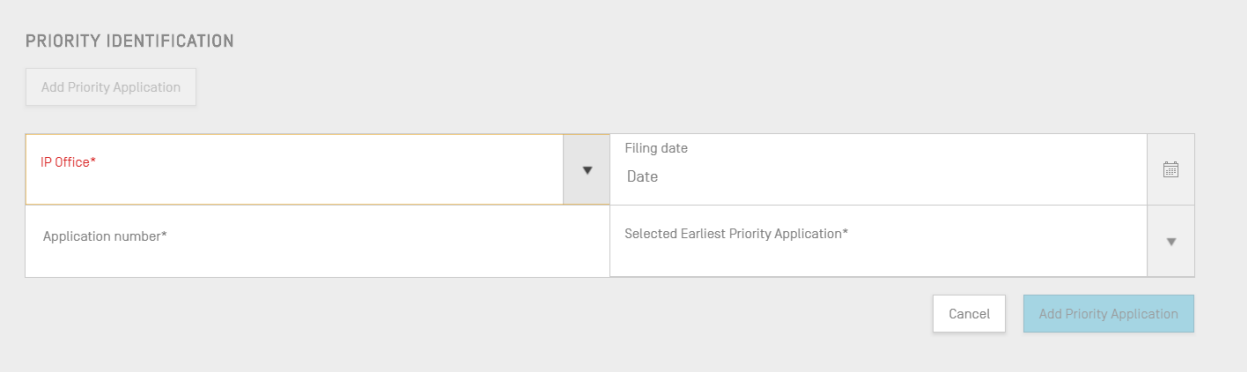


Figure 69 Earliest Priority Application Edit Panel

5.1.2.3 Applicant & Inventor

To add data regarding a new applicant or inventor to the project, the user must click on the “Add Inventor” or “Add Applicant” button within the General Information Section of the Project Detail View. The steps for performing both these actions are identical so only general instructions will be provided but this process must be repeated twice if both an applicant and an inventor are to be included within the project, even if the applicant is also the inventor.

An Overlay will open with two radio buttons, shown in Figure 70. If “Existing applicant/inventor” is selected, the user can choose from a drop-down box which lists currently saved persons and organizations within the local instance of the desktop tool. Figure 70 shows a list of three existing applicants, including “John Smith”.

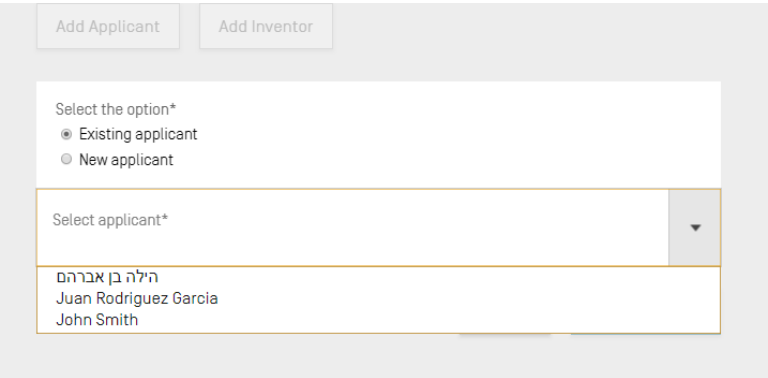


Figure 70 Add existing Applicant/Inventor

If “New applicant/inventor” is selected, the user must fill out the Edit Panel in the same manner as when a new person/organization is being created (see Section **4.2.1** and Figure 71).

Note, only one applicant is required for the sequence listing to be considered valid. As such, one applicant and/or inventor must be marked as primary. This is the applicant/inventor that will appear in the generated sequence listing.

Finally, once the details are complete, the user clicks on the “Add Applicant/Inventor” button shown in Figure 71. As the mandatory fields provided below are not yet included, the user cannot add the applicant and the blue 'Add Applicant' button is greyed out.

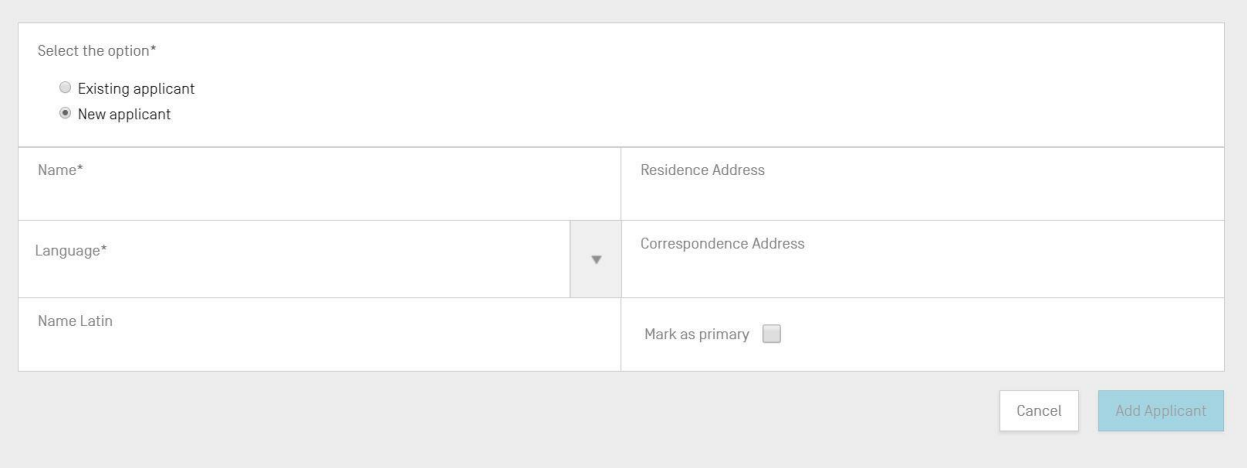


Figure 71 Add new Applicant/Inventor

5.1.2.4 Invention Title

The “Invention Title” is the last subsection within the General Information Section.

1) To add a new invention title, click on the “Add Invention title” button, as shown in Figure 72.

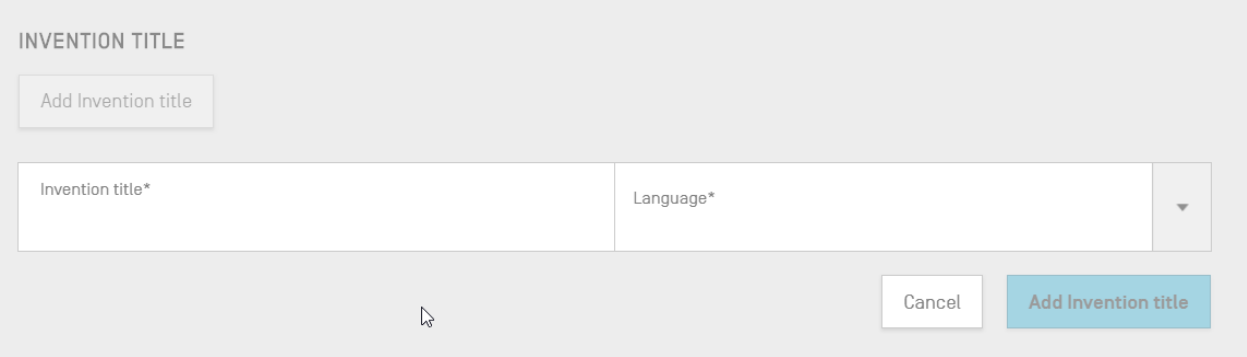


Figure 72 Add Invention title

2) In this Overlay, the user must enter the title of the invention and also indicate what language the title is provided in.

3) Click the blue “Add Invention title” button.

According to the WIPO ST.26, it is mandatory for a sequence listing to have the invention title provided in the language of filing. However a project can also optionally include more than one invention title, in additional languages, but only one invention title per language. Each new invention title can be added using the steps above.

5.1.3 Sequences

The “Sequences” Section of the Project Detail View is where the user provides the technical information related to the sequences themselves.

To create/import/insert/reorder a sequence, the user must scroll down to the ‘Sequences’ section at the bottom of the Project Detail View. The subsections below provide further details on the steps required to perform these actions.

5.1.3.1 Create Sequence

1) Click the “Create new sequence” button, as shown in Figure 73.

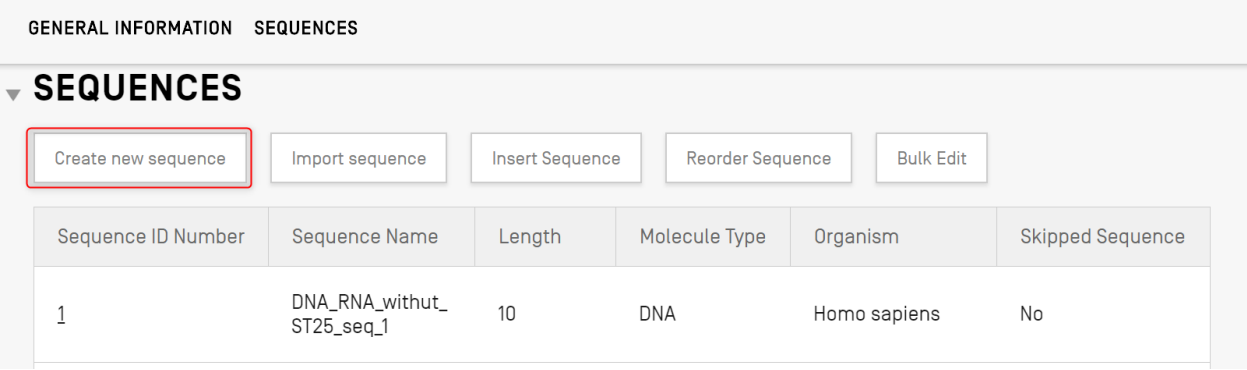


Figure 73 Create new sequence

The remainer of the steps are shown in Figure 74:

2) The user can optionally provide a name by providing a value in the ‘Sequence name’ field, to make it easier to distinguish this sequence. If left blank, the desktop tool will assign a default value with the default name for each new sequence starting with ‘Seq’ and then an iteratively increasing number (‘Seq_1’, ‘Seq_2’, ‘Seq_3’) (1).

The user must select one of the three molecule types allowable in ST.26 and provided by the dropdown box ('DNA', 'RNA' & 'AA') (2). The sequence itself must be entered in the "Residues"[3] text field (3). Note, if the user wishes to create a sequence with both DNA & RNA segments, DNA must be selected as the main molecule type.

The user must also select an organism name for the sequence being created, as this is a mandatory field. This can be either selected from a list of pre-defined organisms in the desktop tool database or one of the custom organisms created and saved locally by the user. For this, the "Existing organism" radio button must be selected (4).

The user also has the option to enter a new (custom) organism name directly by selecting the "New organism" radio button (4).

The user can also classify the organism by selecting a Qualifier Molecule Type[4] from a dropdown box (5) which provides values which vary depending on the Molecule Type previously selected.

If "Mark as an intentionally skipped sequence" (6) is checked, the Sequence panel will remove all constraints on providing values for mandatory elements and the resulting saved sequence will be ignored when validating the project and generating a sequence listing (the residues value will be provided as '000').

If "The sequence contains both DNA & RNA segments" is checked (6), the panel will expand to include fields allowing the user to describe each DNA and RNA segments with a feature "misc_feature". The location of each of the defined segments will be stored in a different "misc_feature" feature (7) of the Sequence, along with a "note" qualifier inside each "misc_feature" with the molecule type followed by the "Further Text" text field value in each "misc_feature". A user can create as many of these features as is necessary by clicking on the "Add new 'misc_feature' feature" button, and this must be done so for all the segments in the whole sequence: both for DNA **and** RNA segments.

[3] See WIPO Standard ST.26 – Annex I, Sections 1 & 3 for the tables of valid Nucleotide symbols & Amino acid symbols

[4] See WIPO Standard ST.26 – Annex I, Sections 6 & 8 for the Qualifier values for the mol type entry under the Nucleotides and Amino acids tables respectively

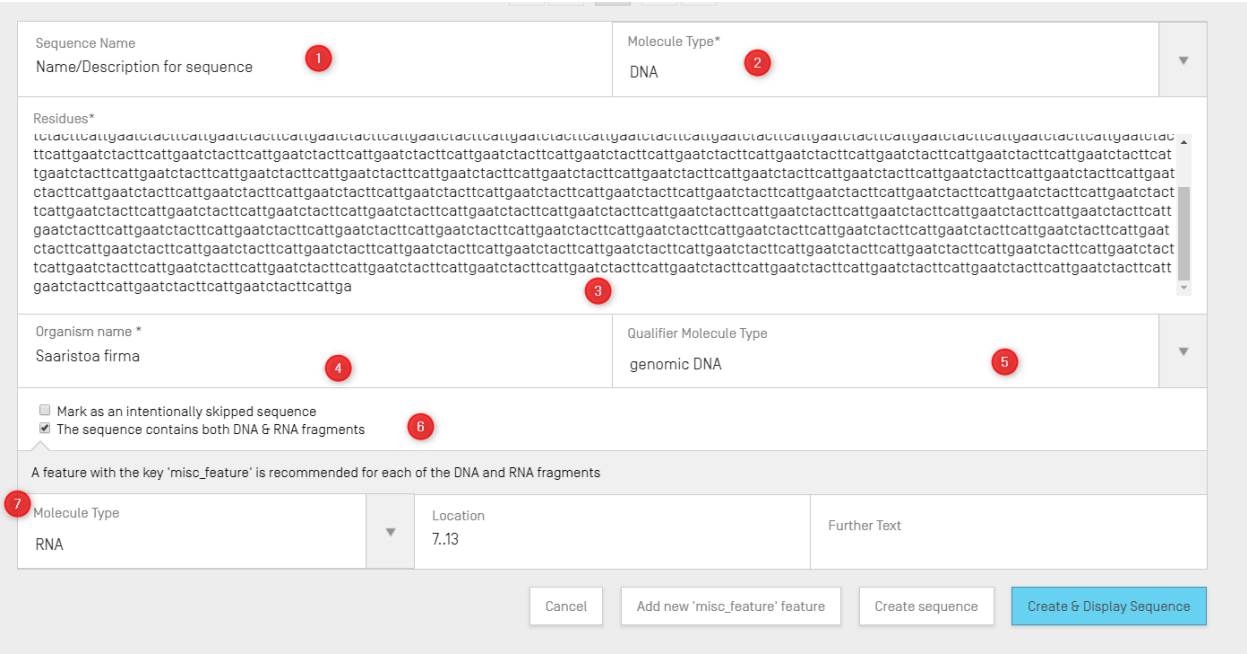


Figure 74 Create new Sequence Panel

3) To finish, the user can click on the grey “Create sequence” button or the blue “Create & Display Sequence” button.

   If the user clicks on the blue “Create & Display sequence” button, a collapsible sequence display will open after creating the sequence, beneath the Sequences Section within the Project Detail View, for the user to review the values.

The newly created sequence can be found in the last position in the list of sequences, with the next available Sequence ID Number.

(To see details on how to reorder the Sequence list, see Section **5.1.3.4**)

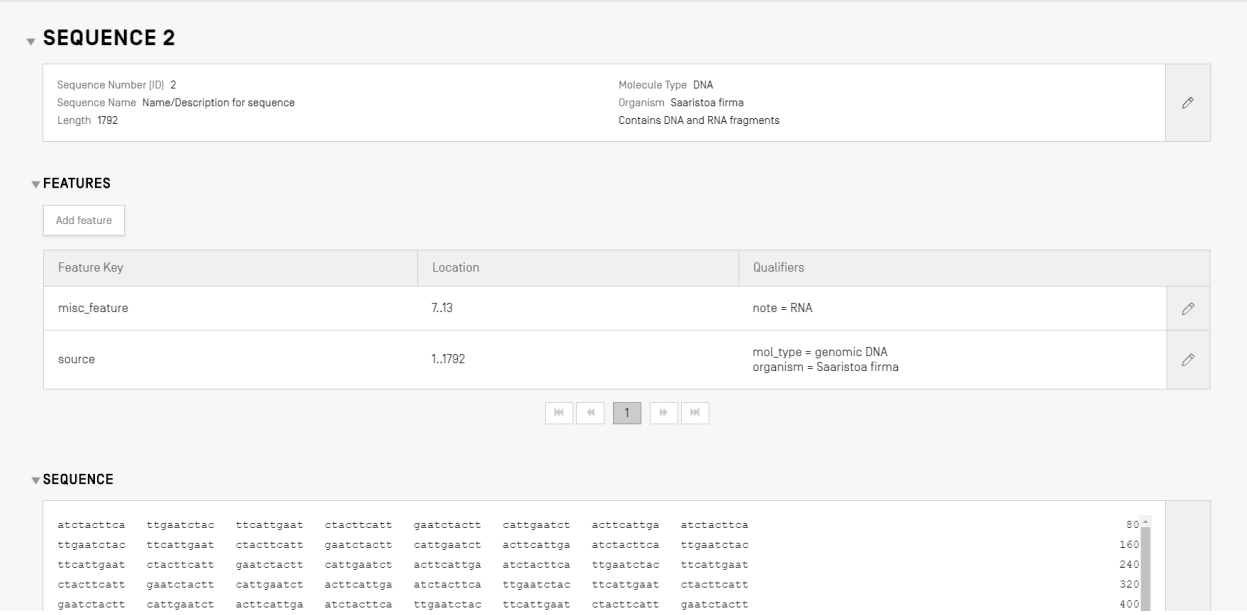


Figure 75 Display sequence

5.1.3.2 Import Sequence

Sequences can also be imported directly from files into a project. The accepted file formats[5] are **raw, multi-sequence, FASTA, ST.26 and ST.25.** When selected, the desktop tool will automatically detect the format used in the file.

1) Click on the “Import sequence button”, highlighted in Figure 76.

[5] See File Formats section of this document (Section 6)

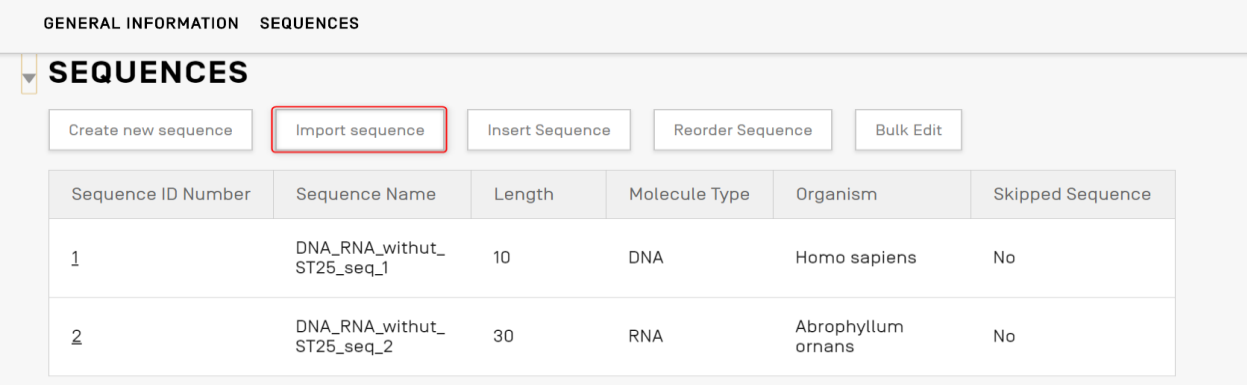


Figure 76 Import Sequence

2) Click on the “Upload file [.txt, .xml]”, shown in Figure 77.

When the dialog box opens, select the file containing sequence data to be imported. The desktop tool will detect the format being used and will perform some validation checks on import.

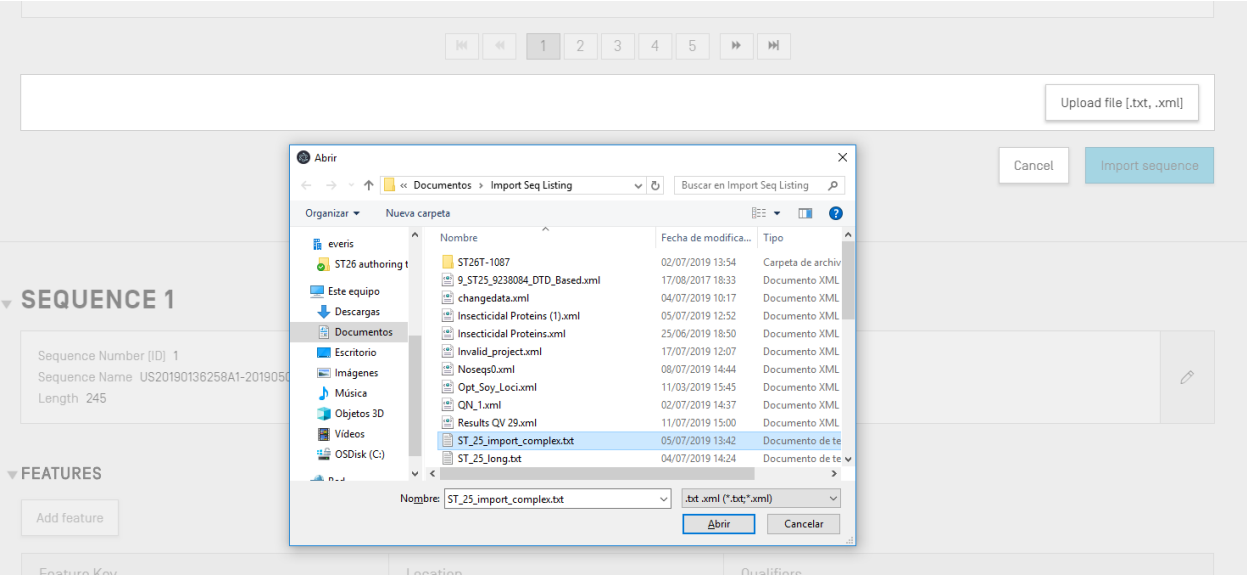


Figure 77 Import Sequence (dialog box)

There are five formats that the tool will accept for importing sequences: raw, multi-sequence, FASTA, ST.26 and ST.25. The format these files must be provided in are outlined in section 6 of this document.

In the case of selecting a file that is in ST.25 or ST.26 format (see Figure 78), the user will first see a “Select Range Sequences” checkbox (1). When checked, this will open a Table (2) with the Sequence ID Numbers of each sequence in the file and the order in which they will be appended to the list of sequence provided in the project.

If the user does not wish to import all the sequences to the project, they can provide the desired range of sequence ID numbers (3).

A single sequence can be entered, as well as a list of sequences separated by commas or a range of sequences in the form `x-y`.

For example: “1, 3, 7, 13-20, 30-50”.

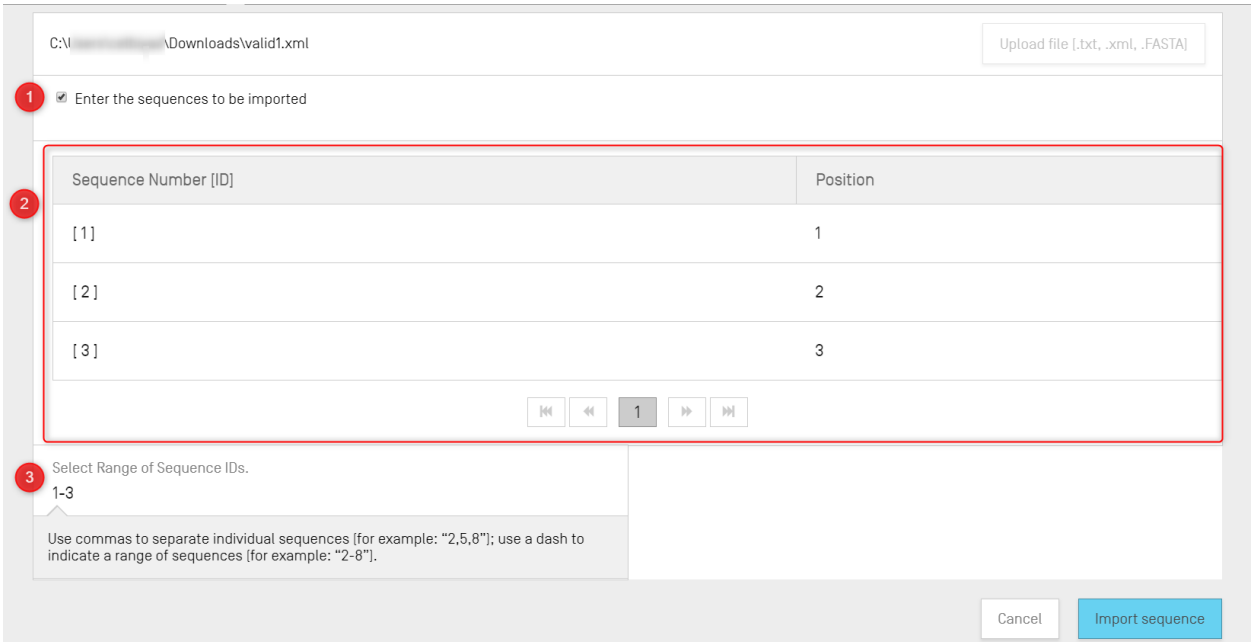


Figure 78 Import Sequence, ST.25 / ST.26 format

In the case of importing a multi-sequence format file (see Section **6.1** for information on this format and Figure 79), the user will see a “Select Range Sequences” checkbox (1), which when checked, will display a preview Table showing the Sequence ID Numbers of the corresponding sequences in the file as well as the details of each sequence under the “Detail” column (2): including sequence name, molecule type and organism name.

The user must select the range of Sequence ID numbers that they wish to import to the list of sequences within the project (3). By default, the total number of sequences of the selected sequence listing file will be displayed as a range.

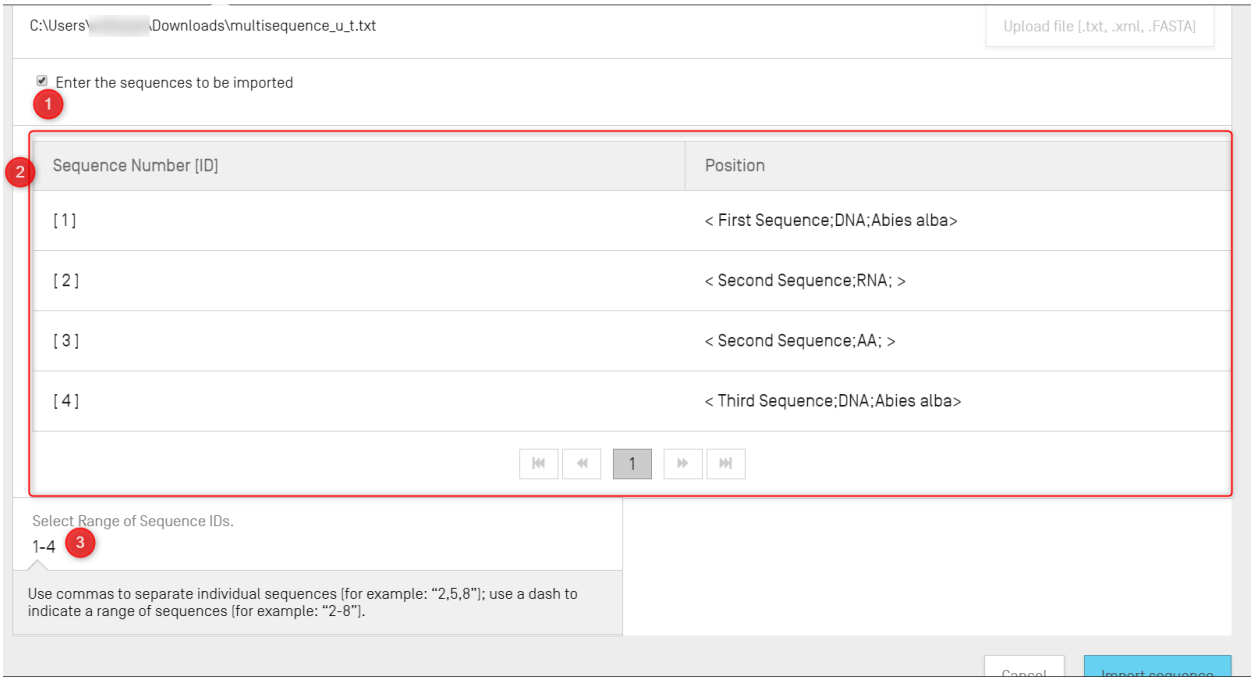


Figure 79 Import Sequence, multi-sequence format

The last case for formats that are accepted by the import Sequence process is the raw file format (see Section **6.3**). This format only defines a single sequence per file. When a raw file is selected for import, the tool will display the Edit Panel shown in Figure 80 and when a FASTA file is imported Edit panel shown in Figure 81 is displayed. The user should proceed by providing the mandatory fields.

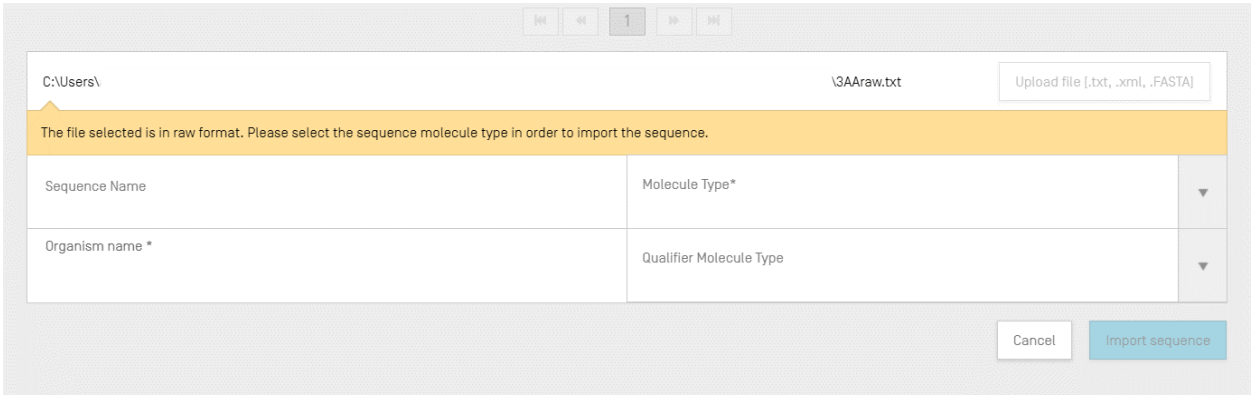


Figure 80 Import Sequence, RAW format

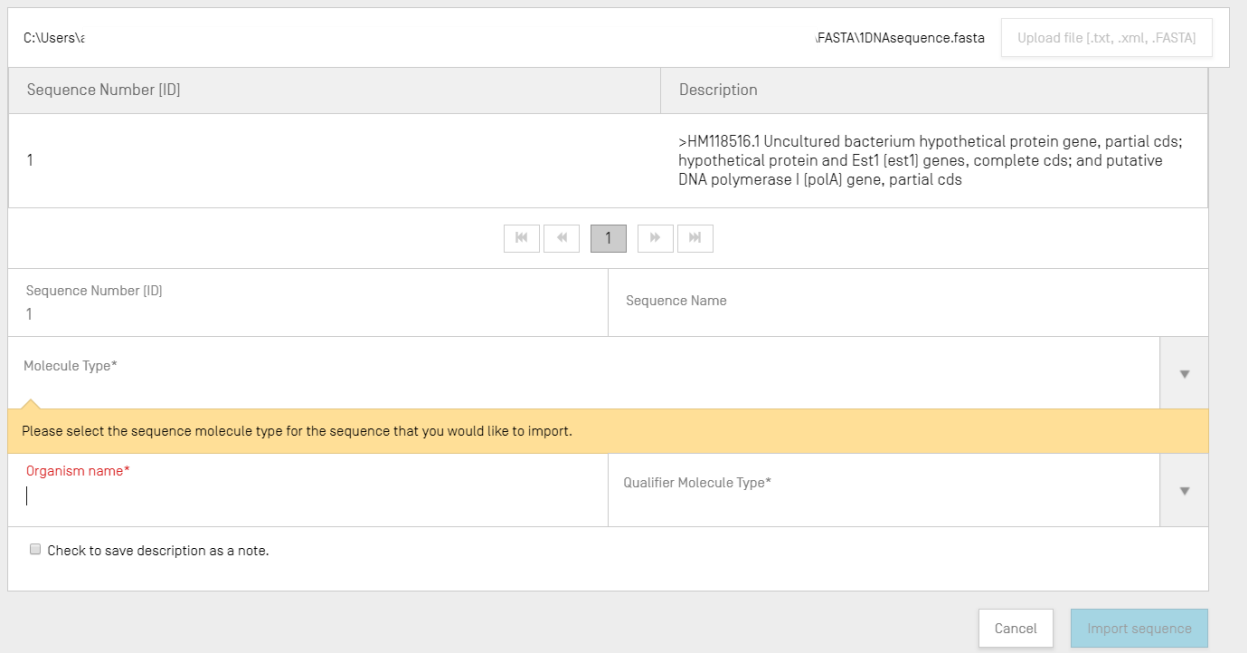


Figure 81 Import Sequence, RAW format

3) To finish, the user should click on the “Import sequence” blue button.

After the import, the tool will navigate to the “Import Report” View, detailed in Section **5.4.**

5.1.3.3 Insert Sequence

To insert a sequence into a specific position of the list of sequences, the user must click on the “Insert Sequence” button at the top of the Sequences Section (highlighted in Figure 82).

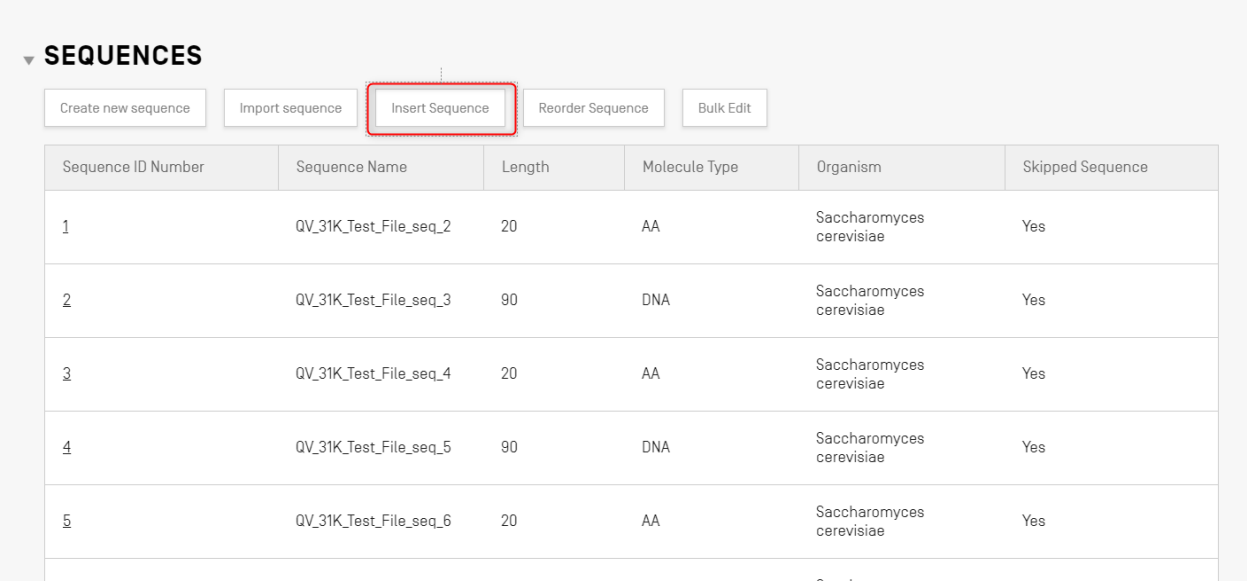


Figure 82 Insert sequence

An Overlay with a panel will appear (see Figure 83). The user must fill out all the information required for creating a sequence (see Section 5.1.3.1), and in addition, at the top-left of the panel, the user must enter the position in which the sequence should appear in the list of sequences (highlighted in Figure 83).

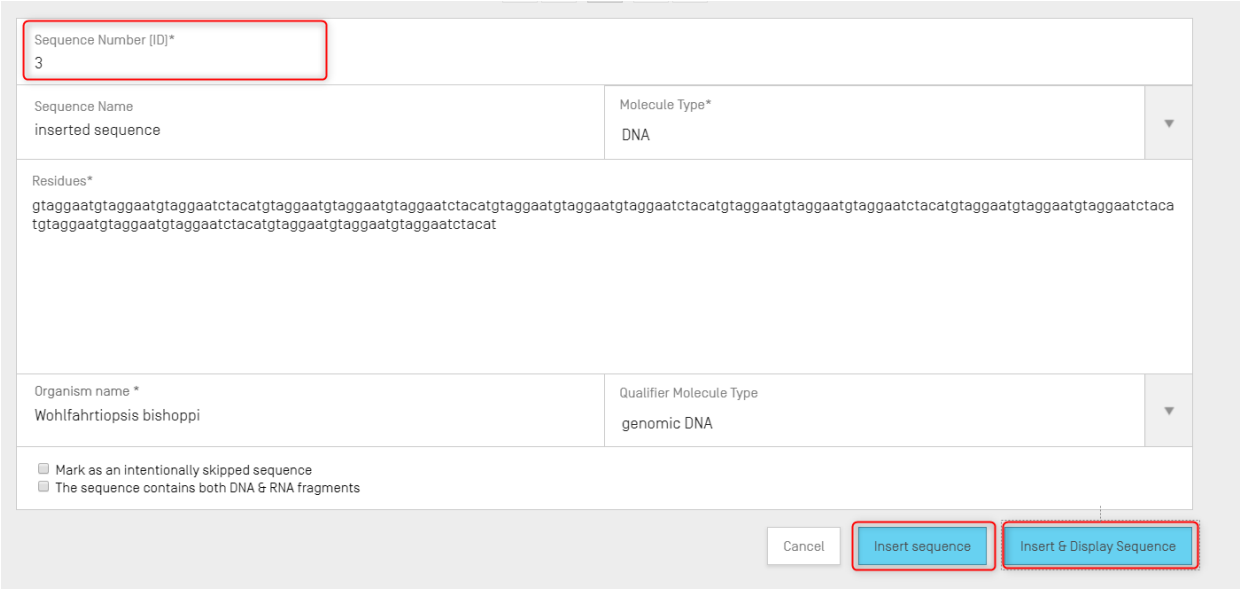


Figure 83 Insert sequence panel

To finish, the user can click on the insert “Insert sequence” or “Insert & Display Sequence”, also highlighted in Figure 83.

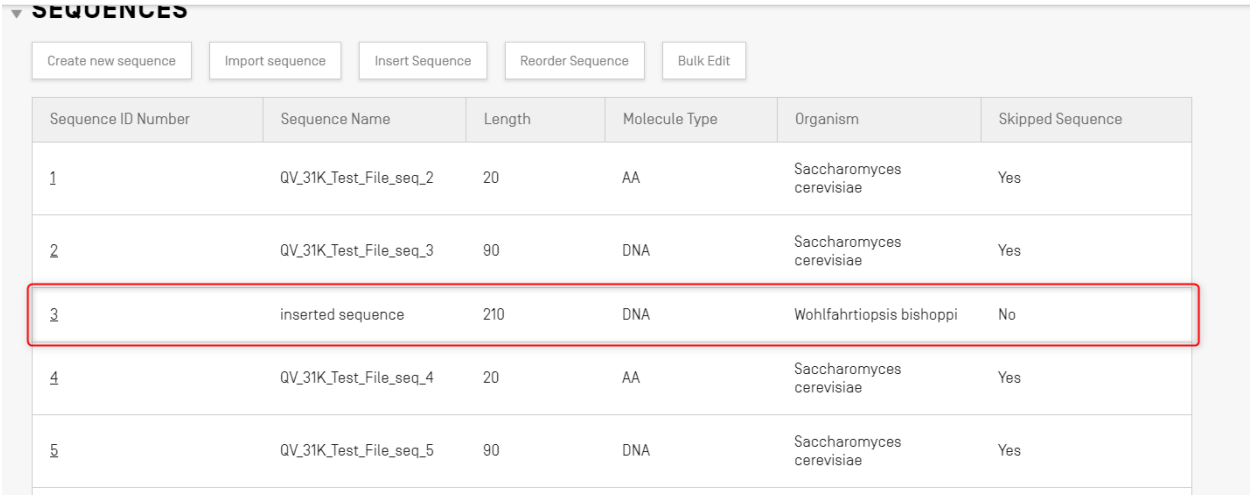


Figure 84 Sequence Inserted

If the user clicks on the blue “Insert & Display sequence” button, a collapsible sequence Table will open after creating the sequence, beneath the list of sequences within the Project Detail view, as shown in Figure 84.

5.1.3.4 Reorder Sequence

The user can reorganize in what order the sequences should appear within the list of sequences provided in a project by using the following steps. The sequence order transformation is shown from Figure 85 (start) to Figure 87 (result).

1) Click on the “Reorder Sequence” button, shown in Figure 85.

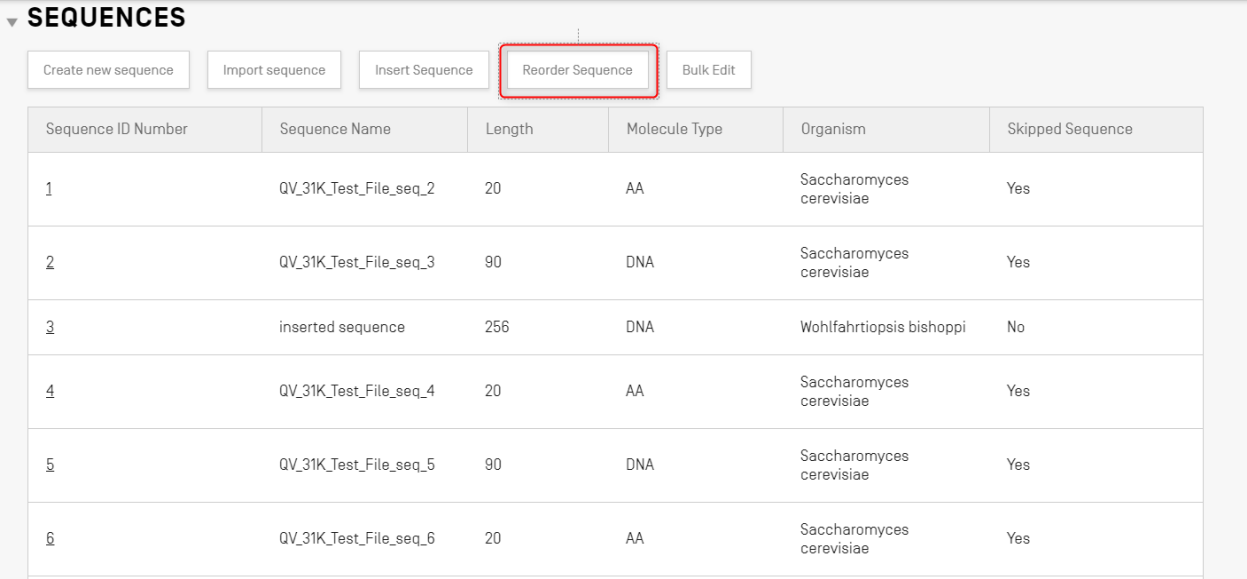


Figure 85 Reorder Sequence

2) Select the sequence(s) to be moved and select the position in which they are to be placed with respect to the current sequence list order, as shown in Figure 86.

Figure 86 Reorder Sequence panel

3) The sequence now appears in Figure 87 in the second position.

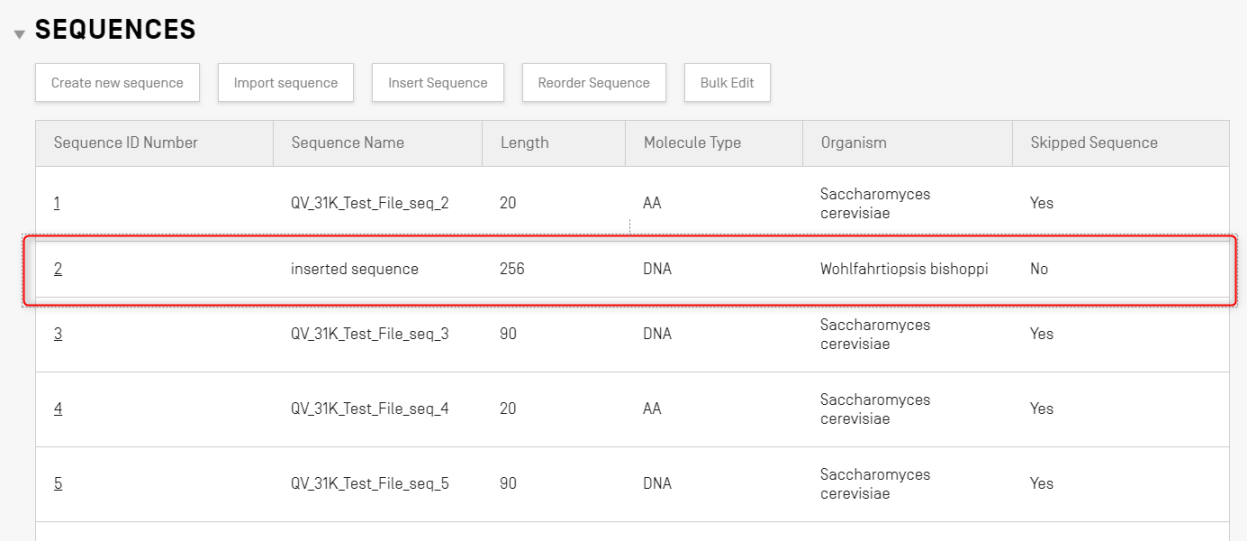


Figure 87 Reordered Sequence

5.1.3.5 Bulk Edit

The user can use Bulk Edit when changes need to be made to multiple sequences. While you can go into sequences individually and edit, this would be unfeasible for projects with a large number of sequences.

1) Click on “Bulk edit”, shown in figure below:

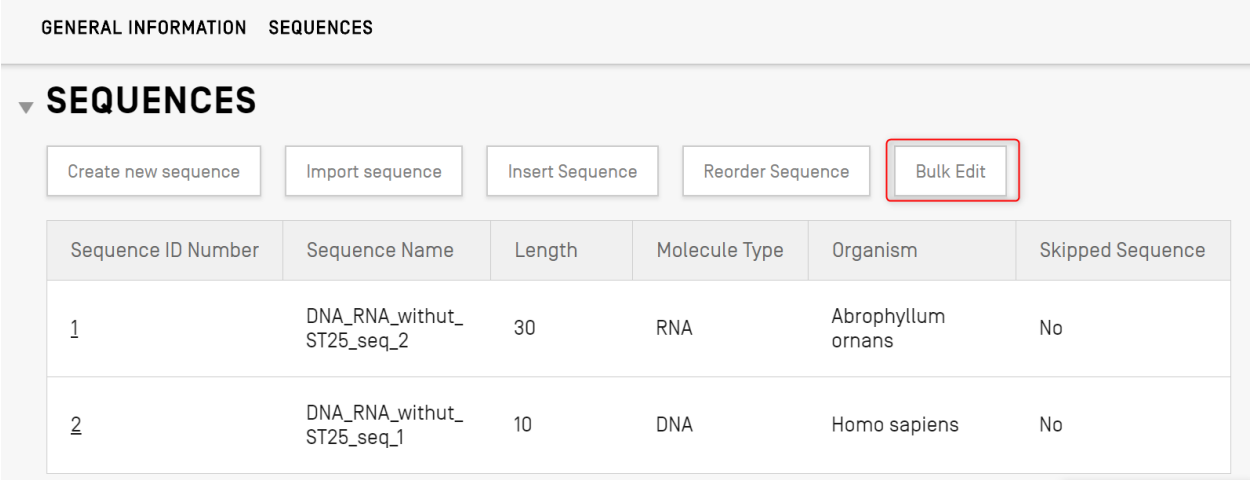


Figure 88 Bulk Edit

2) Choose “Type of bulk edit” as 'Qualifier molecule type', 'Organism' or 'Feature':

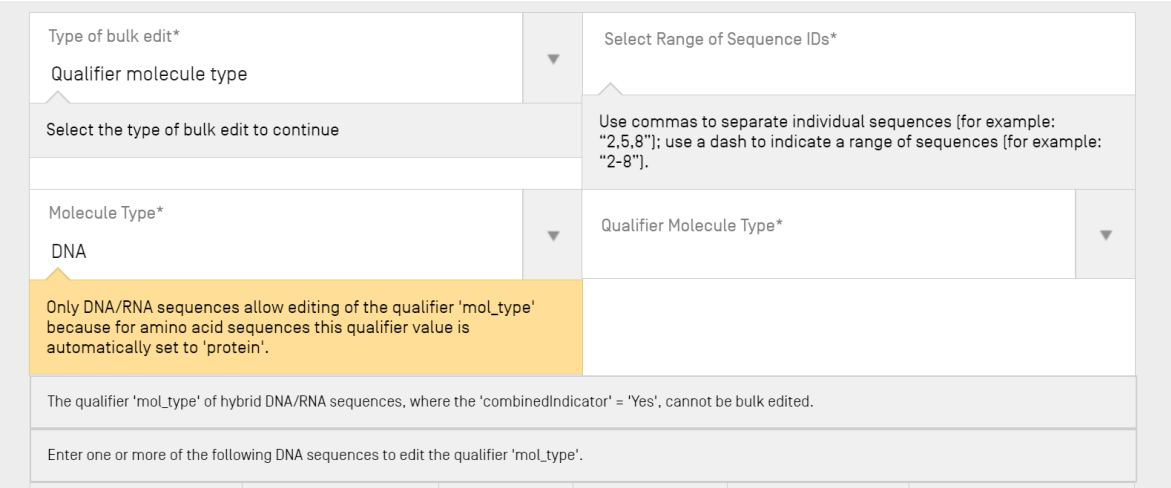


Figure 89 Type of bulk edit

As shown in Figure 89 , the system informs the user that ONLY nucleic acids sequences can have the value of the qualifier ‘mol_type’ edited (because the same value for the amino acids sequences is automatically set by the system to ‘protein’).

3) After selecting the Molecule Type, the system prompts the user to select the type of nucleic acids sequences to which the bulk edit will apply

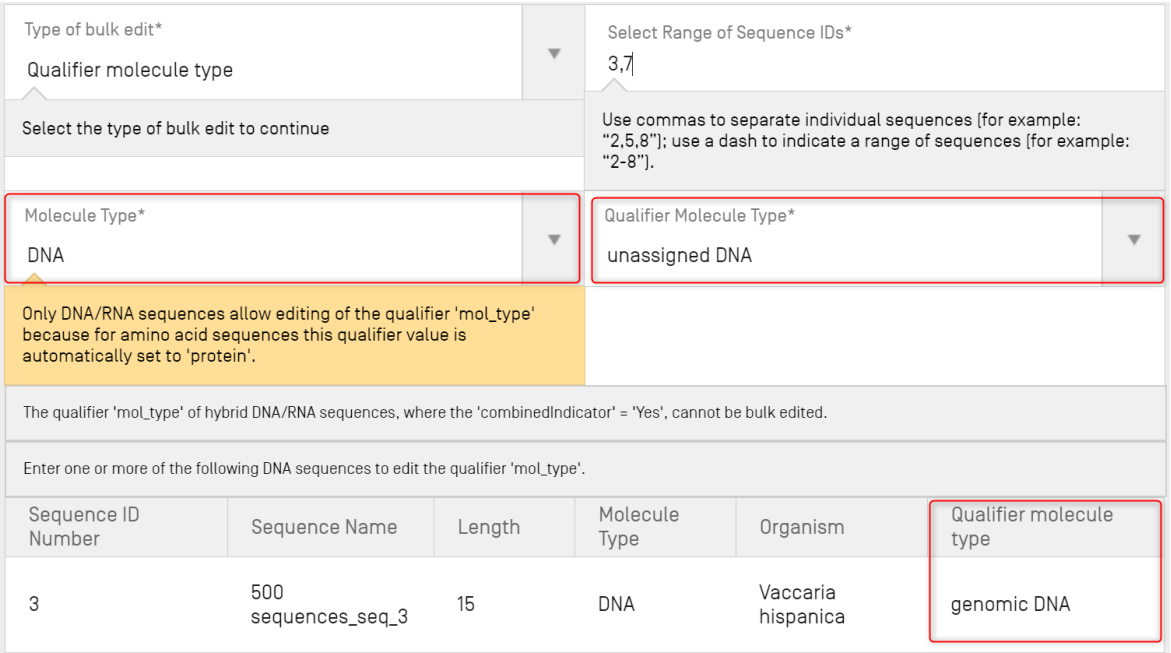


Figure 90 Qualifier Molecule Type to update

As shown in Figure 90, the system also warn users that the qualifier “mol_type” for sequences where organism = “synthetic sequence” must be “other DNA” or “other RNA”, and if they change these values, an error will be generated on project validation.

4) The system displays for selection the list of sequences to be bulk edited

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

Figure 91 Updated Qualifier molecule type

When the type of bulk edit is Organism:

1) Choose “Type of bulk edit” as Organism (1). The user must enter the range of sequence IDS to be edited (2). Then if the user has chosen to modify the value of organism by “synthetic construct”, the system will notify him that the Qualifier molecule type will be automatically changed to “other DNA” or “other RNA” according to molecule type. (3)

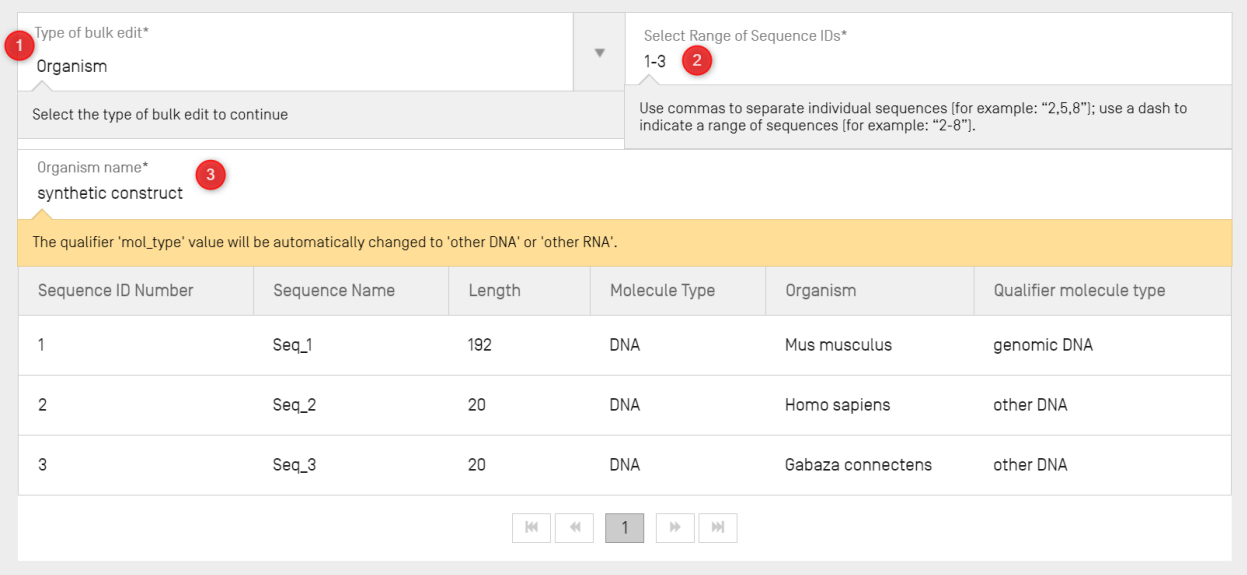


Figure 92 Bulk edit - Organism

When the type of Bulk edit is Feature:

1) Choose “Type of bulk edit” as Feature (1). The user must enter the range of sequence IDS to be edited (2) and the molecule Type (3). Then if the user has chosen to modify the value of the feature location by “complement(join(1..30,61..90))”, the tool will update all existing CDS feature for sequence 2 and 3 by the provided value.

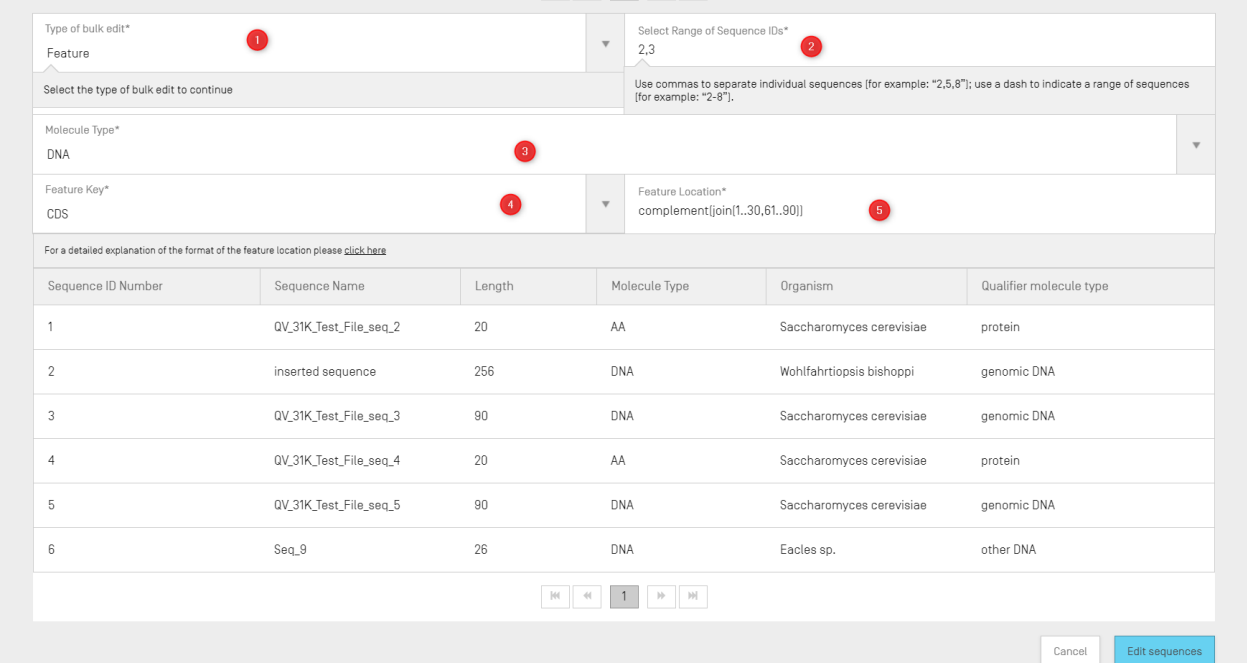


Figure 93 Bulk edit - Feature

5.1.3.6 Edit Sequence

To edit a sequence, the user must click on the Sequence ID Number of the corresponding sequence to be modified, as highlighted in Figure 94 (Sequence ID ‘1’).

Note:

All sequence listings compliant with WIPO Standard ST.26 must start with SEQ ID #1 and be numbered consecutively until sequence length.

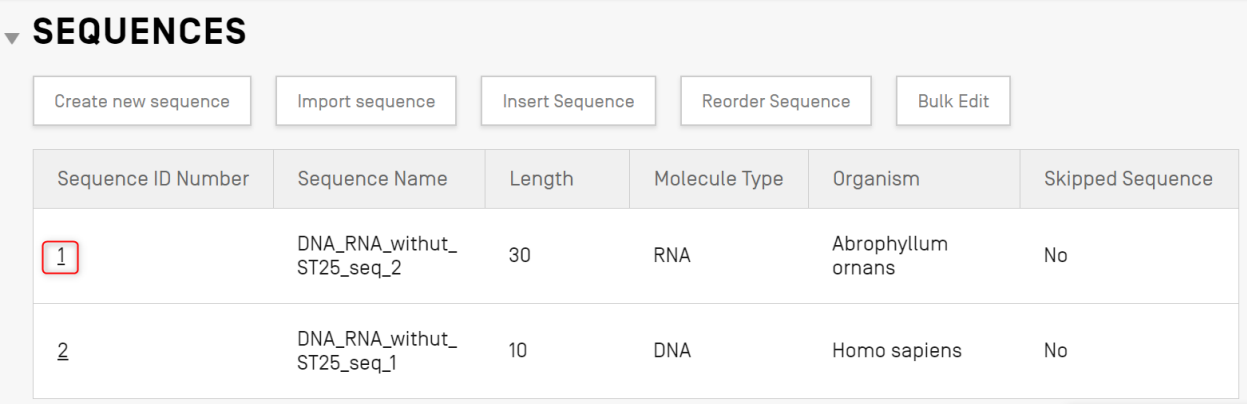


Figure 94 Edit sequence

This will open a new Section in the Project Detail View beneath the Sequences Section, as shown in Figure 95.

This Sequence Section is composed of the “Basic Information” Edit Panel (1), the “Features” list Table (2) and the “Sequence” residues Edit Panel (3).

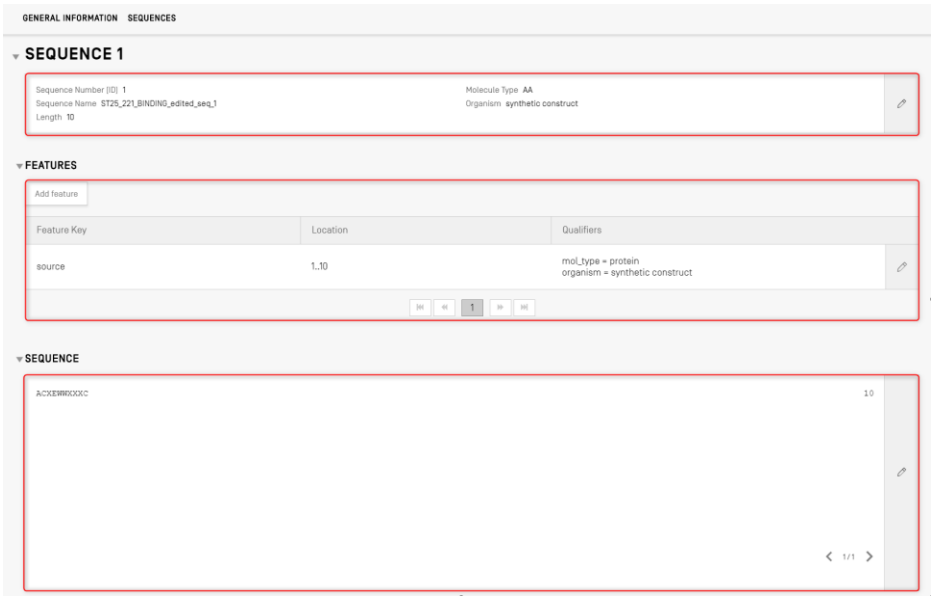


Figure 95 Edit sequence, display sequence

5.1.3.7 Features

According to WIPO ST.26, every sequence MUST have at least one feature associated with it: 'source', depending on the molecule type. Each source feature must have two mandatory qualifiers: organism and mol_type.

The Features Table has three columns: the feature key, the location of the feature within the genetic sequence and the qualifiers associated with an individual sequence feature.

The feature location indicates in which segment of the sequence the feature exists. The allowable formats to specify the feature location are provided in WIPO ST.26 and are as follows:

- Single residue number: *x*
- Residue numbers delimiting a sequence span: *x..y*
- Residues before the first or beyond the last specified residue number: *<x, >x, <x..y, x..>y, <x..>y*
- A site between two adjoining nucleotides: *x^y*
- Residue numbers joined by an intrachain cross-link: *x..y*

Location **operators** can be used to form complex location descriptions:

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

To add a new feature to the sequence, click the "Add feature" button in the Features Section of the selected Sequence, highlighted below in Figure 96.

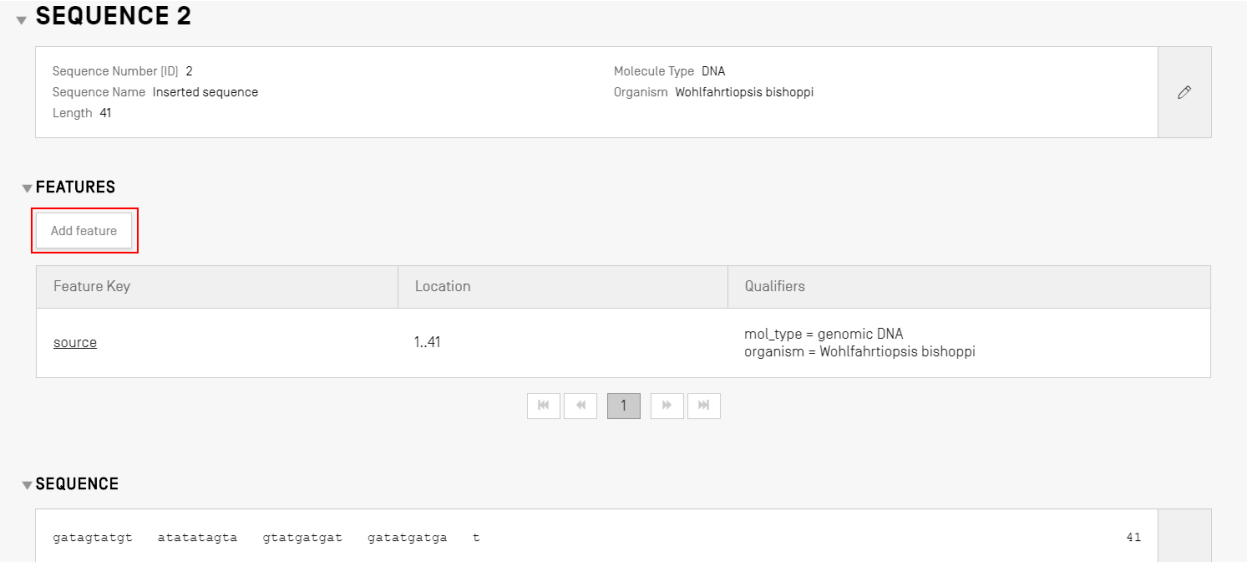


Figure 96 Add feature

Next, in the Overlay that opens (shown in Figure 97) select an entry from the feature key[6] dropdown list and specify the feature location within the sequence that the feature applies to.

[6] See WIPO Standard ST.26 – Annex I, Sections 5 & 7 for tables of feature keys for Nucleotides and Amino acids

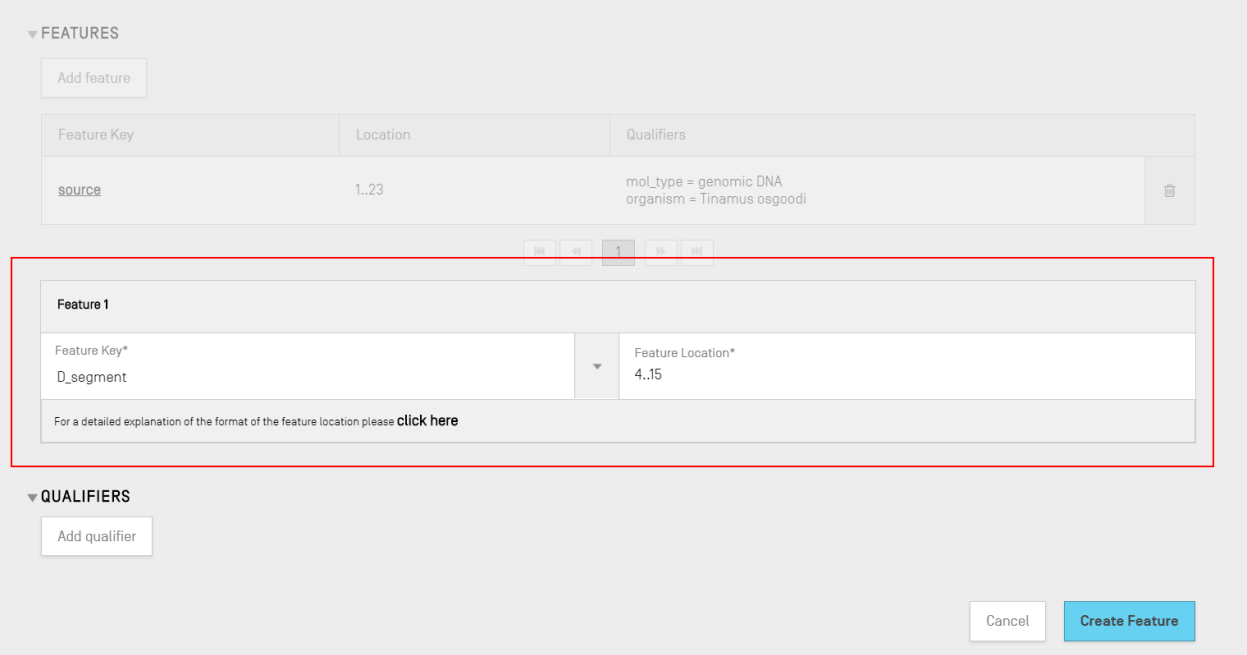
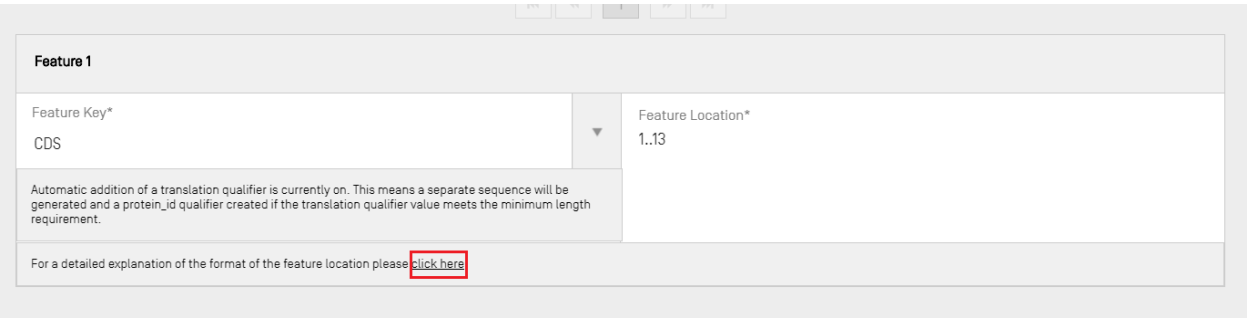


Figure 97 Add feature, overlay

Qualifiers can also be added to the feature at this stage, but they will be covered in the next section (Section **0**).



A link has been added to have a detailed explanation of the format of the feature location. The link is located at the end of the sentence ‘click here’. The link has been underlined in Figure 98 to make it easier for the user to identify it.

Figure 98 Feature Location Link

To finish, click the blue “Create Feature” button, shown in Figure 97.

5.1.3.7.1 CDS Features

The CDS Feature type is used to describe the coding sequence for a protein[7]. A CDS feature may optionally include the amino acid translation of the segment of the sequence to which it belongs and if this satisfied the minimum length requirement will appear as a separate sequence within the project. Within the CDS Feature of the original sequence, there is a reference to the Sequence ID of the translated amino acid sequence provided in the “protein_id” qualifier.

When creating a “CDS” feature for a sequence, the ‘translation’ qualifier (with default “Genetic Code” value of 1 – “Standard Code”) can be automatically added to the CDS feature with a qualifier value of the translated a residue chunk of the sequence as indicated by the feature location. An associated 'protein id' and separate amino acid sequence may also be generated by checking the checkbox in Basic Information provided at the top of the project details page (see Figure 102). However, this qualifier is not mandatory and can be deleted after generation. The user can also manually create a 'translation' and 'protein_id' qualifiers which references the associated translated Sequence ID which has also been created by the user.

Automatic CDS Feature creation

The steps for automatically creating a CDS feature qualifier are as follows:

1) In the specific sequence display, click the “Add feature” button and select “CDS” as the feature key. If the checkbox ‘automatically add a translation qualifier’ in Basic Information is checked, it will automatically add a translation qualifier, its value, and a protein_id qualifier and its associated separate amino acid sequence (if appropriate) when a CDS feature is added to a nucleotide sequence.

2) The user also has the option to manually create a translation qualifier.

[7] https://www.ddbj.nig.ac.jp/ddbj/cds-e.html

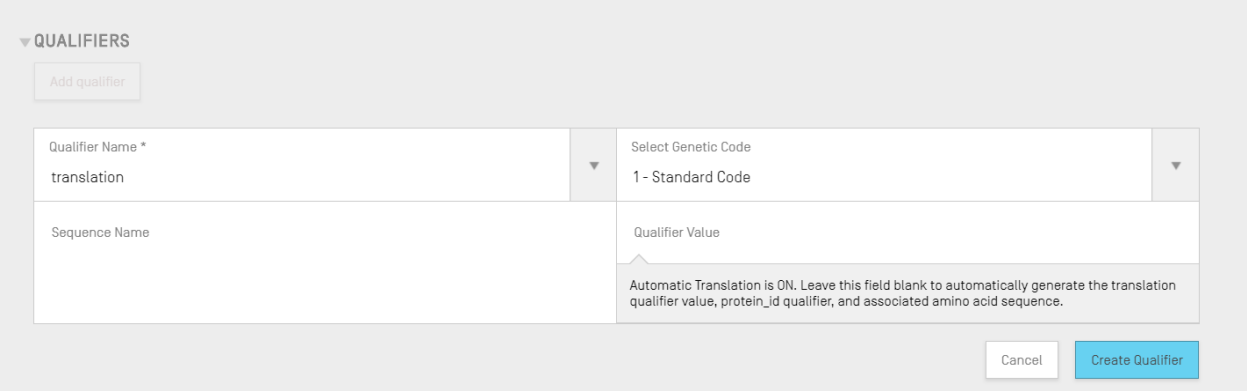


Figure 99 Create translation qualifier

3) When the user is finished editing the feature and its related qualifiers, they must click the "Create Feature" button to save it. A resulting CDS feature is shown in Figure 100**Error! Reference source not found.**.

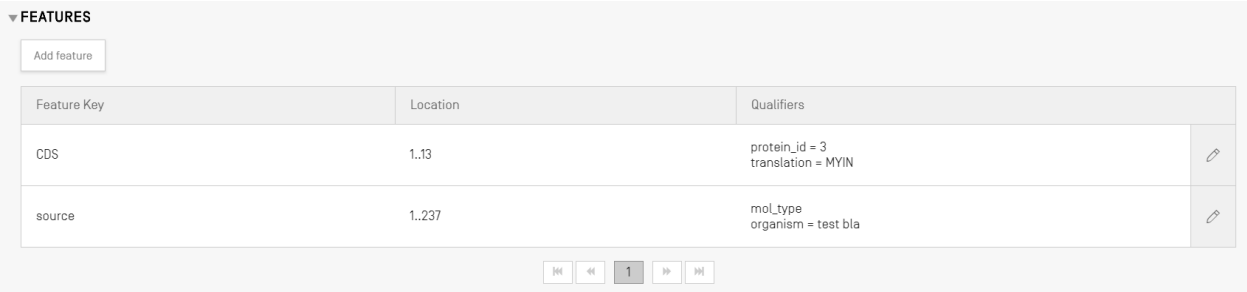


Figure 100 CDS feature

The tool then ***creates a new sequence**** for the project with the following attributes:

- **Sequence ID Number** = the next available value for Sequence ID Number
- **Length** = length of the translated sequence
- **Sequence Name** = the value given in the "Sequence Name" field of the "translation" qualifier. If no name was provided, the default sequence name will be provided ('Seq_#').
- **Molecule Type** = "AA"
- **Organism Name** = the same value as provided for the original sequence
- **Qualifier Molecule Type** = "protein"
- **Sequence Residues** = translated values of the original sequence

A resulting example after this process has completed is shown in Figure 101.

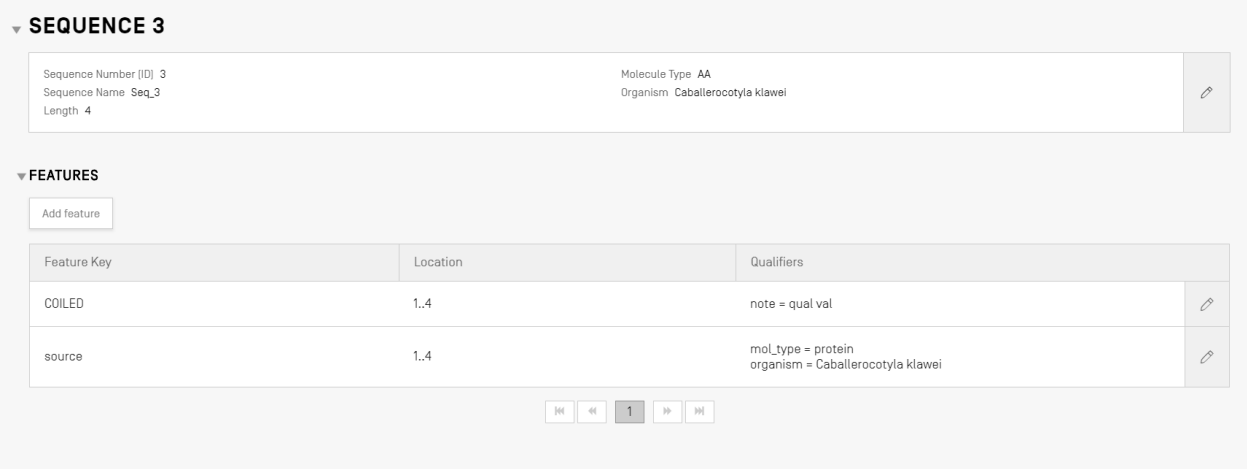


Figure 101 CDS feature, generated sequence

Note, regarding the creation of the translated sequence:

The separate translated sequence is created only if it has least 4 specifically defined residues defined, (e.g.,AXTG counts as 3 characters).

In the case of modifying the "translation" qualifier, if the qualifier value includes less than 4 specifically defined residues, then the associated sequence translation will be removed, as will the 'protein_id' qualifier.

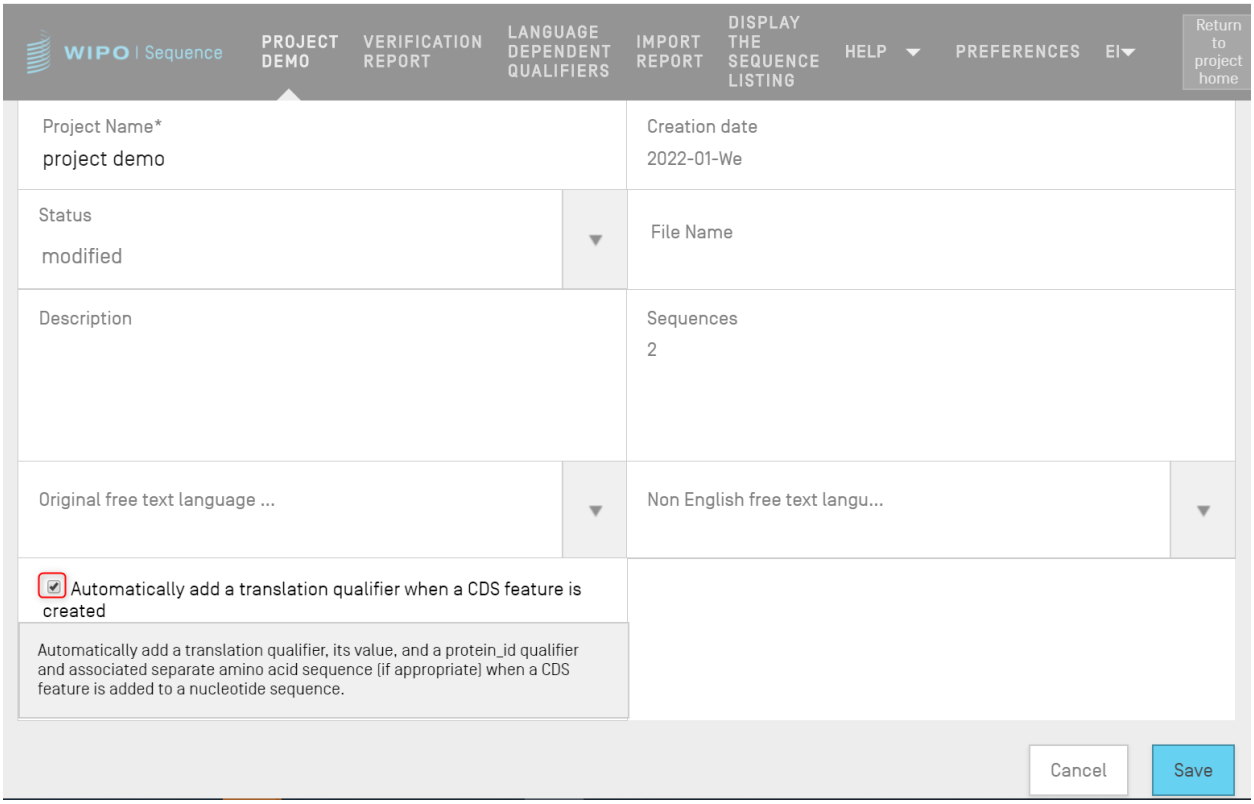


Figure 102 Automatic CDS translation treatment : check to turn on

Advice around CDS features when including a pseudo or pseudogene qualifier:

Make sure auto-translation is turned off when adding a pseudo or pseudogene qualifier to a CDS feature.

If auto-translation is not turned off when a pseudo or pseudogene qualifier is added to a CDS feature, then when the CDS feature is updated, a translation qualifier will automatically be added. To correct this error, turn off auto-translate for the project, then open the CDS feature and delete the translation and protein_id qualifiers, and then update the feature.

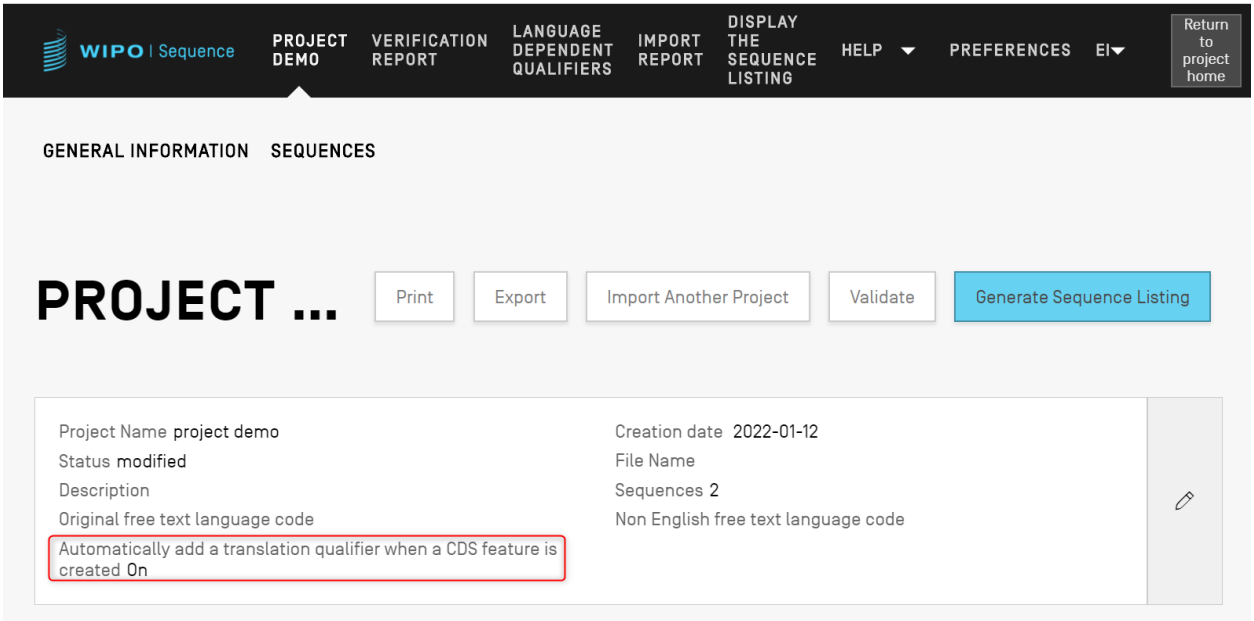


Figure 103 Automatic CDS translation treatment is On

If the user wishes to automatically generate the translation qualifier, the translation table value and sequence name can be set from the Edit Panel of the qualifier. When the user creates the feature, the tool will perform the translation and then add a “protein_id” qualifier to the feature and a new sequence with the value of the translation.

The translation will be performed again, only if the feature location or one of the qualifiers “transl_table”, “transl_except”, or “codon_start” changes its values, in which case the linked sequence will be updated.

***Note**:*

If the translation value is changed, the linked sequence will update its value automatically. However, if the linked nucleotide sequence is modified, the value of the translation qualifier will not change. If the “protein_id” qualifier is modified after creation, then the linked sequence will lose its association to the original sequence.

Advice around use of stop codon:

Typically, stop codons should only be found at the end of a CDS feature, indicating the end point of the encoded amino acid sequence. They should never be found in the middle of a CDS feature unless there is a transl_except qualifier that indicates that the stop codon is to be translated into a particular amino acid.

If a stop codon is found in the middle of a CDS feature (highlighted in yellow below), and there is no transl_except qualifier indicating that the stop codon is to be translated into a particular amino acid, as shown in Figure 104**Error! Reference source not found.**, then the tool should stop

translation at that point and a red banner would be displayed informing the user that no translation will be generated.

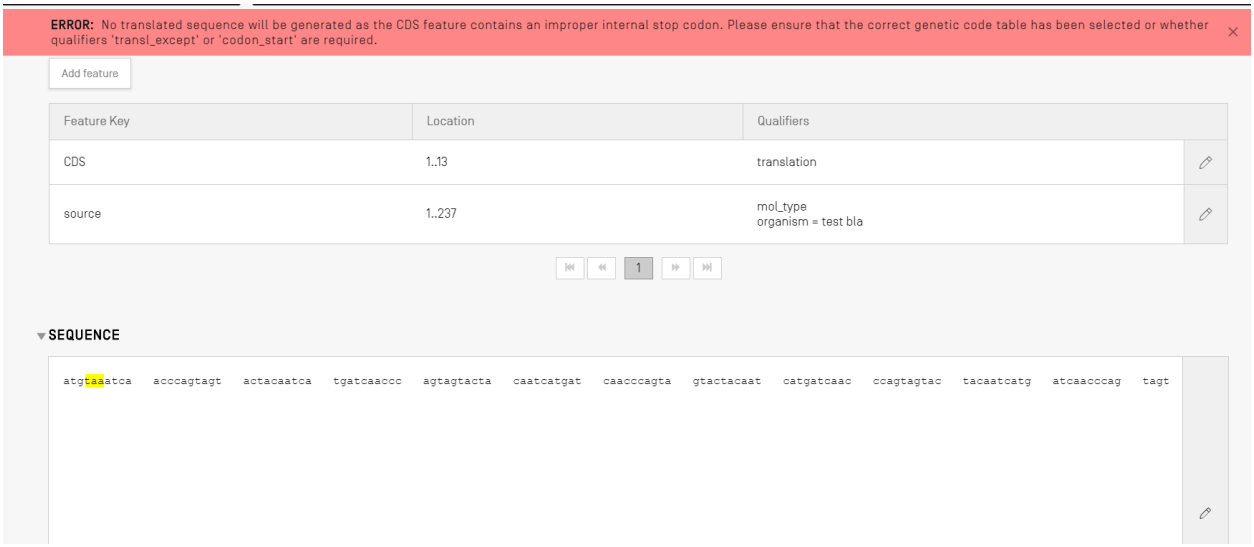


Figure 104 Automatic translation checkbox is On and the sequence contains an improper stop codon

And an error should be listed in the validation report to alert the user that there is a problem with their coding sequence, as shown in Figure 105.

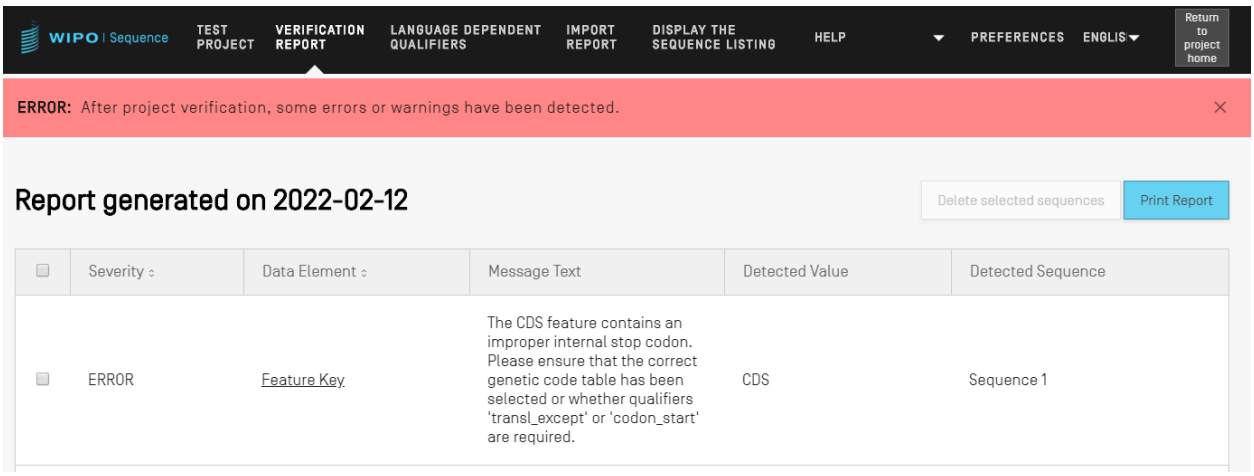


Figure 105 sequence contains an improper stop codon verification error

5.1.3.8 Qualifiers

To view the qualifiers for a feature, the user must first select the relevant feature from the Feature Table of the relevant sequence. In the example shown in Figure 106, the pencil icon for the 'source' feature is highlighted.

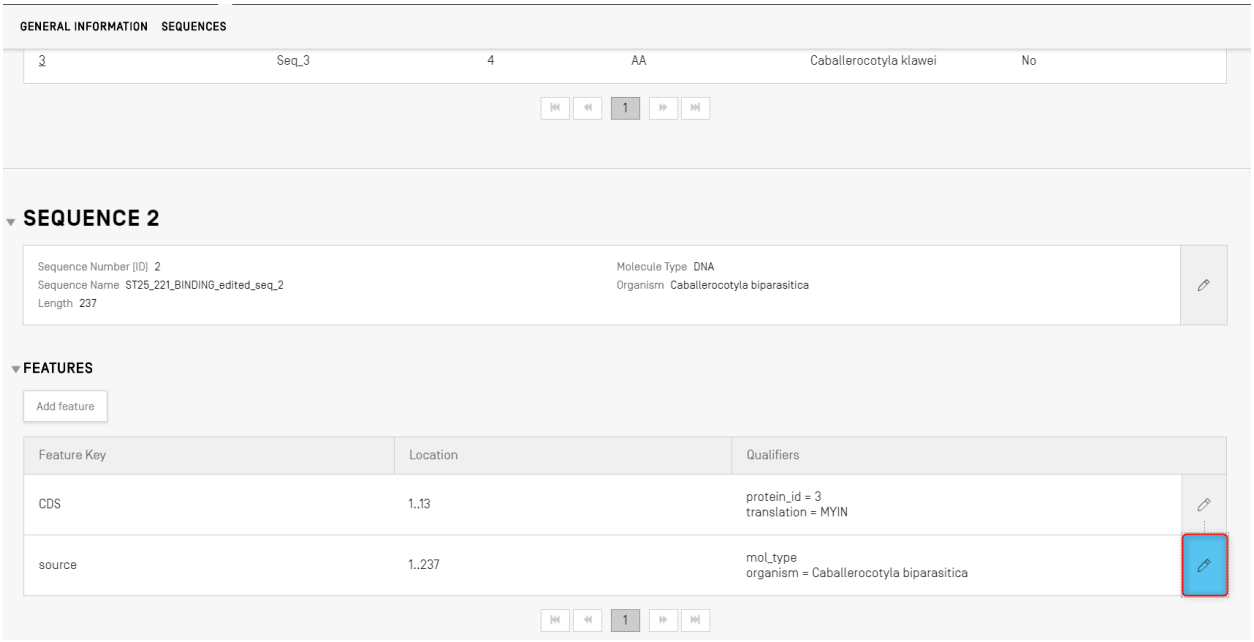


Figure 106 View qualifiers

Clicking on the pencil button will open the following feature Overlay, as shown in Figure 107.

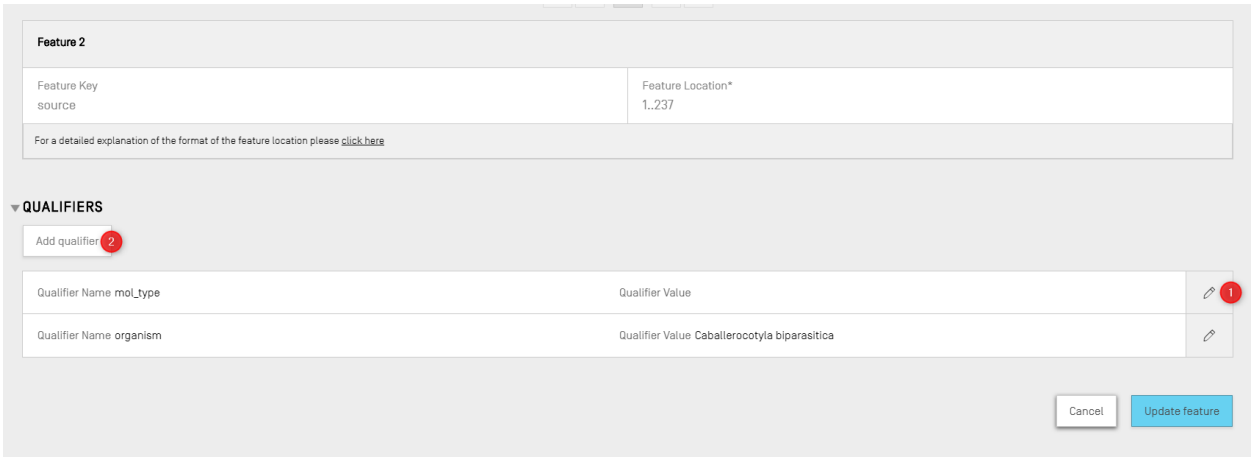


Figure 107 Add/Edit qualifiers

Existing qualifiers can be edited by clicking on the pencil icon to the right of each row (1), or the user can add a new qualifier to the currently selected feature by clicking the “Add qualifier” button (2).

When editing or adding a qualifier, the user will be presented with the two fields shown in Figure 108: the ‘Qualifier name’[8] (to be selected from a dropdown list) and the ‘Qualifier value’.

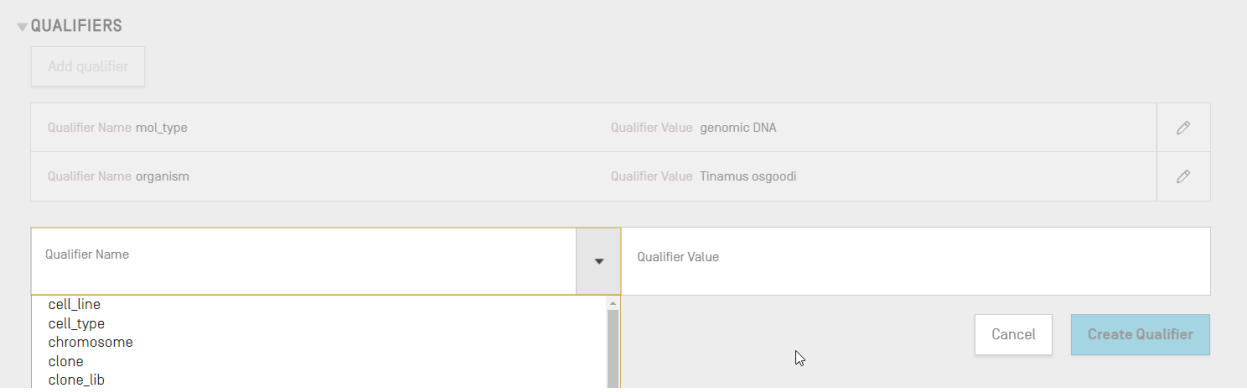


Figure 108 Qualifier Panel

The Qualifier Value field will have a different behaviour depending on the type of qualifier:

- **Qualifiers with pre-defined values**. The value field is a dropdown field where the user can select one of the predefined values for the qualifier, as shown in the example below:

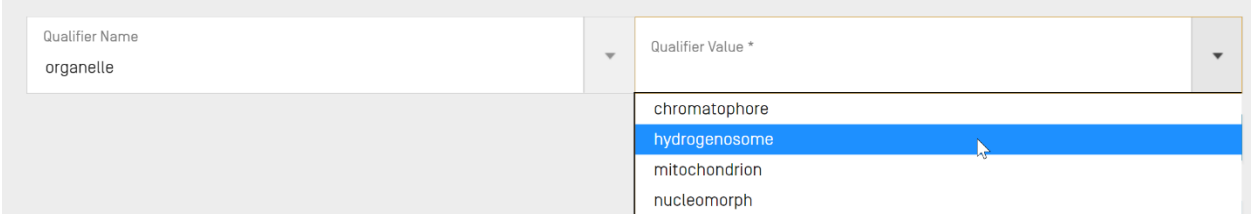


Figure 109 Qualifiers with pre-defined values

[8] See *WIPO Standard ST.26 – Annex I, Sections 6 & 8* for all possible Qualifiers for Nucleotides and Amino acids

- **Qualifiers with free text**. The value field is a free-text field. In addition to the Qualifier Name and the Qualifier Value, which holds the English value only, two additional fields appear to allow the user to provide both the language code (e.g., 'ru') and the corresponding language value in the Non English Qualifier Value, as shown in the example below:

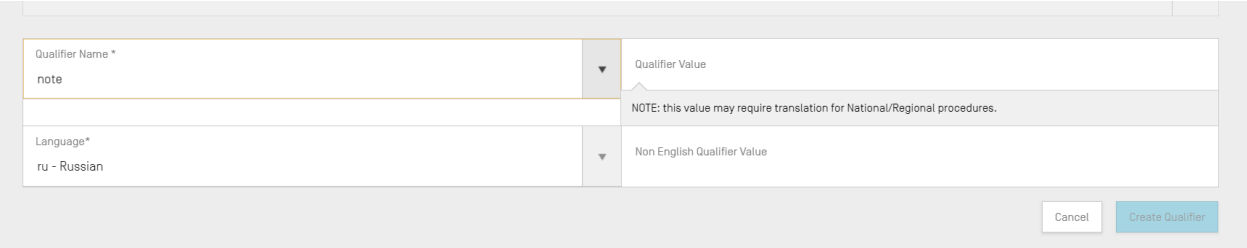


Figure 110 Qualifiers with free text

The Language code field is assigned the same value as the 'Non English Free text language code' filed in the Project Detail Information.

The user can provide a series of Non English values for each selected language either by manual input or by importing the proper associated language from an XLIFF File (see section **5.1.3.2** for more details on Import).

- **Qualifiers with pre-defined format**. The value field is a free-text field, but the value entered is validation to ensure it matches the specific rules provided in WIPO ST.26 Annex I, Section 6[9]. In the example shown, the date has not been provided in the correct format:

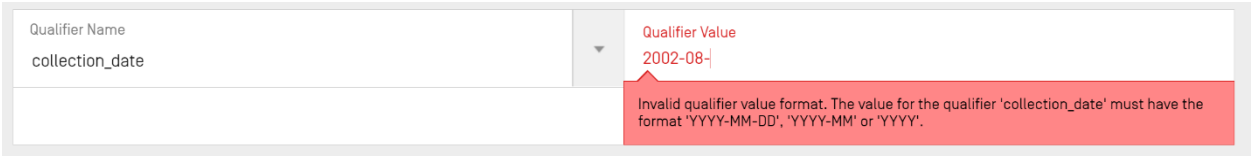


Figure 111 Qualifiers with pre-defined format

[9] See WIPO Standard ST.26 – Annex I, Section 6 for rules regarding these qualifiers.

- **Qualifiers with no value allowed**. The qualifier value field is not editable, as indicated below in the example:

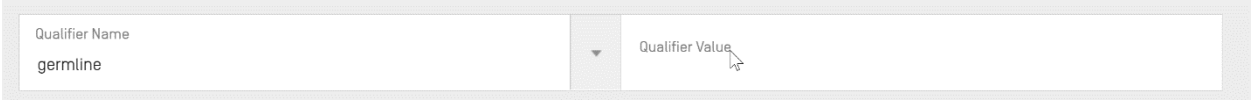


Figure 112 Qualifiers with no value allowed

Note:

The feature location for the source feature cannot be edited after it is created and the mol_type and organism qualifiers can only be edited and not deleted.

When finished, the user must click the blue “Create Qualifier” button to add the newly created qualifier, or “Save”, to save the changes made to the existing qualifier.

The last step, once the qualifier(s) have been added/modified, the user must click on the “Update feature” button at the bottom of the Feature Overlay, shown in Figure 113 to proceed.

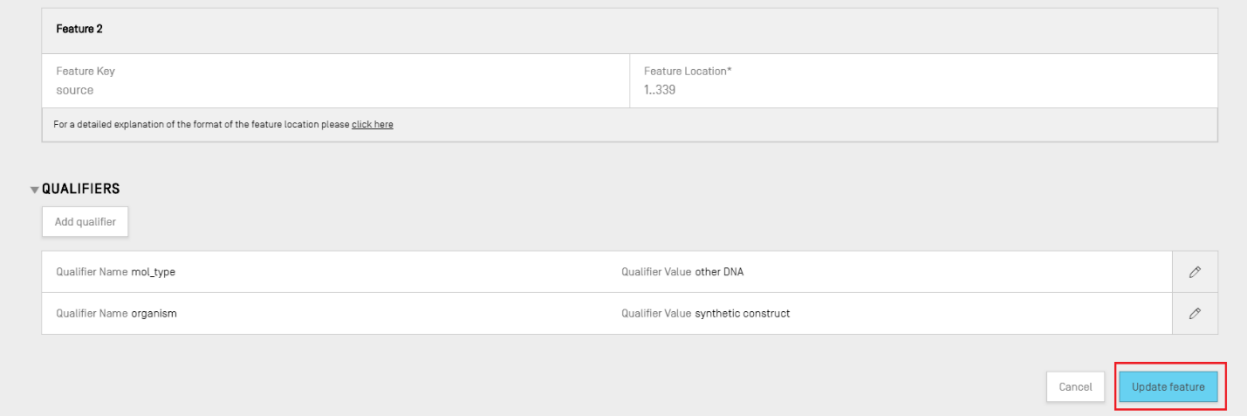


Figure 113 Update feature

5.2 VERIFICATION REPORT

To open the verification report of the sequence listing of a project, from the Project Details View, the user can click on the “VERIFICATION REPORT” link in the menu bar at the top of the View.

For further details on how to generate the verification report, see Section **5.1.1.3.**

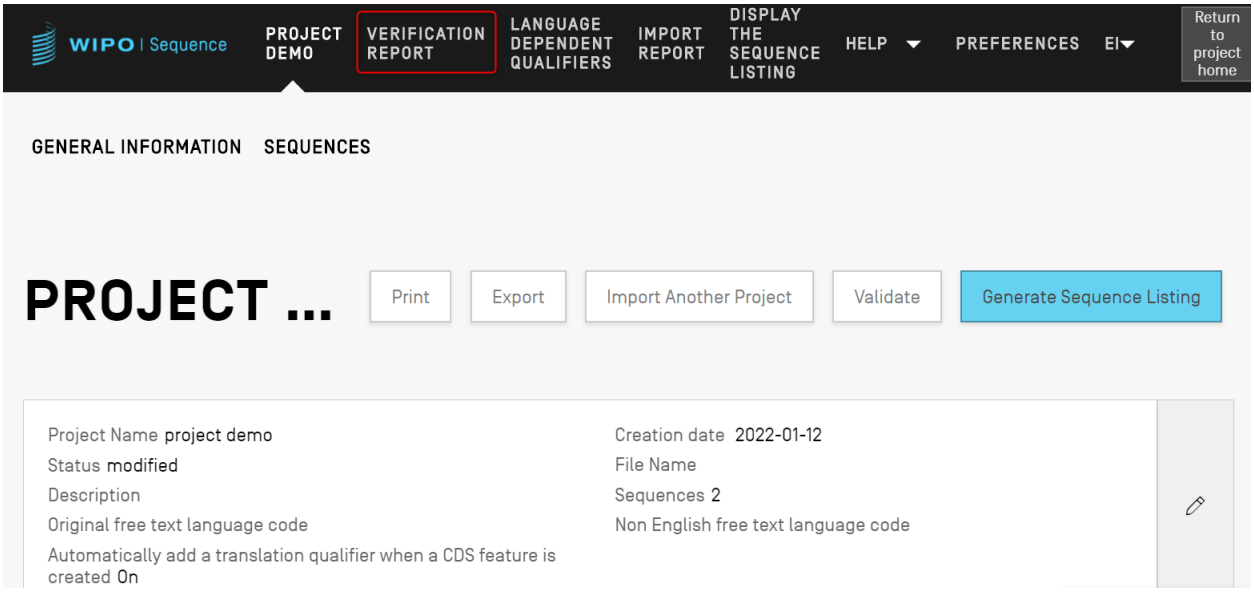


Figure 114 Verification view/button

Depending on whether the project sequence listing is valid or not there will be two resulting Views: Figure 115, where errors/warnings have been generated and Figure 116, after a successful validation.

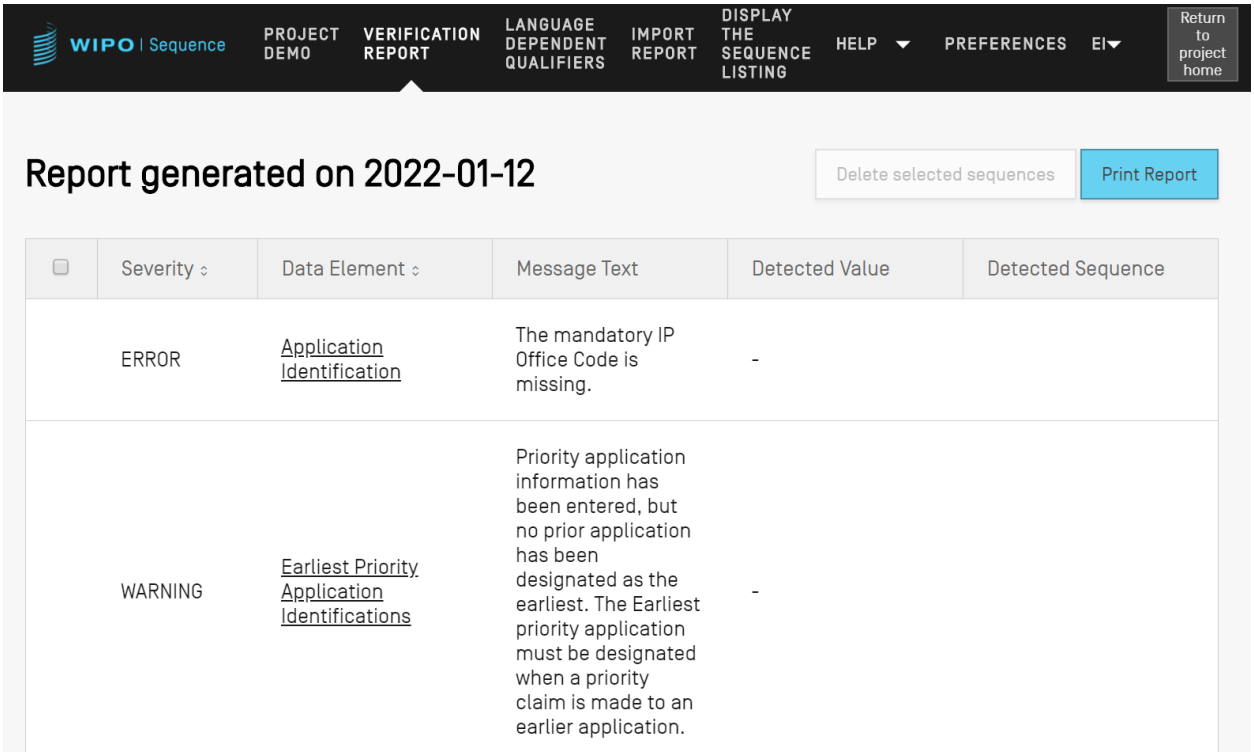


Figure 115 Verification report, errors

The user has the option to bulk delete sequences by checking the checkbox in the first column of the verification report table and clicking the button "DELETE SELECTED SEQUENCES" (see Figure 115).

The Verification Report can be exported as a PDF by clicking on the “Print Report” button, shown in the top-right-hand corner of Figure 115. The generated report will be displayed in the PDF viewer as explained in Section **3.2**.

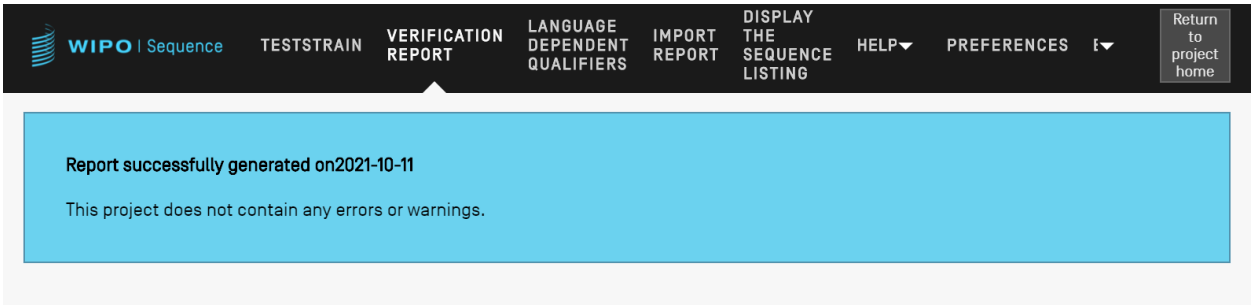


Figure 116 Verification report, no errors

5.3 LANGUAGE DEPENDENT QUALIFIERS

The qualifiers which allow a “free text” value in a project are further referenced within the “LANGUAGE DEPENDENT QUALIFIERS” view of the Project page.

Whenever a language dependent qualifier is added to the current project, the qualifier will also be displayed in this View, shown in Figure 117.

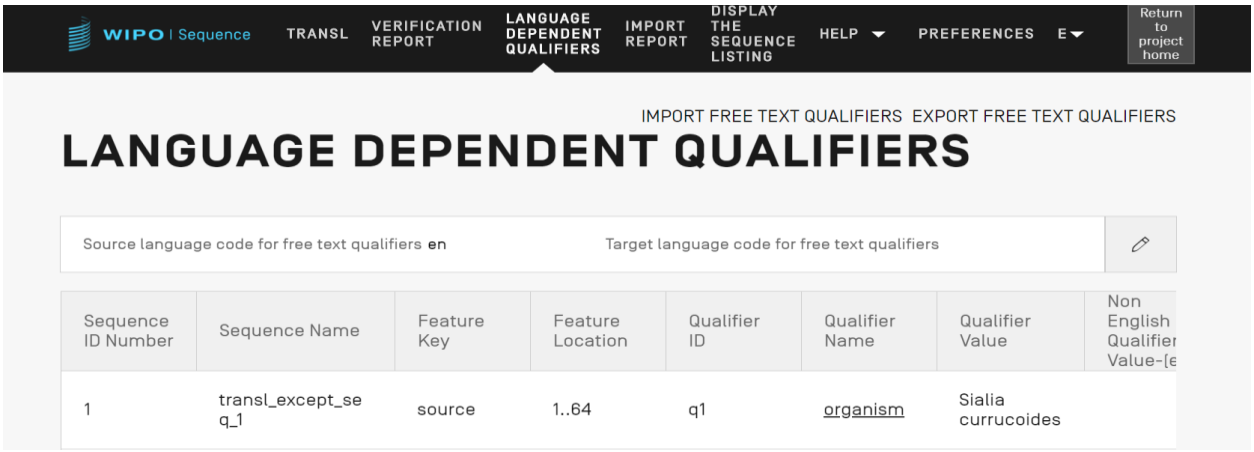


Figure 117 Language dependent qualifiers view

The user can modify a qualifier’s associated translated free-text value by clicking on the ‘Qualifier Name’ value, as shown in Figure 107, which will open an Overlay with an Edit Panel underneath the table.

The user will need to provide the source language code and target language code for free text qualifiers XLIFF file export where the translated values will need to be provided by translators before reimporting the XLIFF file.

5.3.1 IMPORT FREE TEXT QUALIFIERS

If the user clicks on the “IMPORT FREE TEXT QUALIFIERS” button, the tool will open the file explorer so the user can browse to find and select the (. XLIFF) file to import. Multiple validation steps are provided to ensure that the correct mappings between the source and target language values are conducted.

The selected file must contain the following items of data:

- Project name
- The target language code
- The source language code

- For each XLIFF unit element:
    - The qualifier unique ID (following the format: a number preceded by the letter ‘q’)
    - The qualifier value in the source language tag
    - The qualifier value in the target language tag

Figure 118 An example of a valid XLIFF File

Once the user has confirmed the selected file for import, the tool will ask the user to verify if they want to proceed by confirming a series of verification steps (see Figure 119 to Figure 123).

- The system compares the project name from the input file with the name of the selected project:

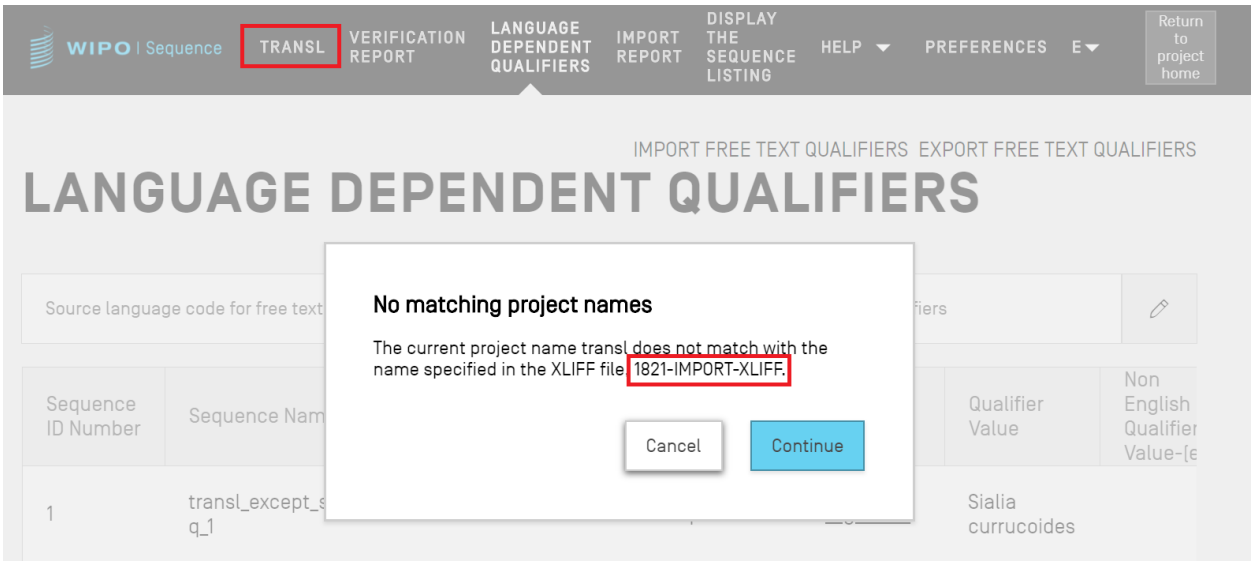


Figure 119: Project name validation

- The system will inform the user if any qualifiers could not be mapped:

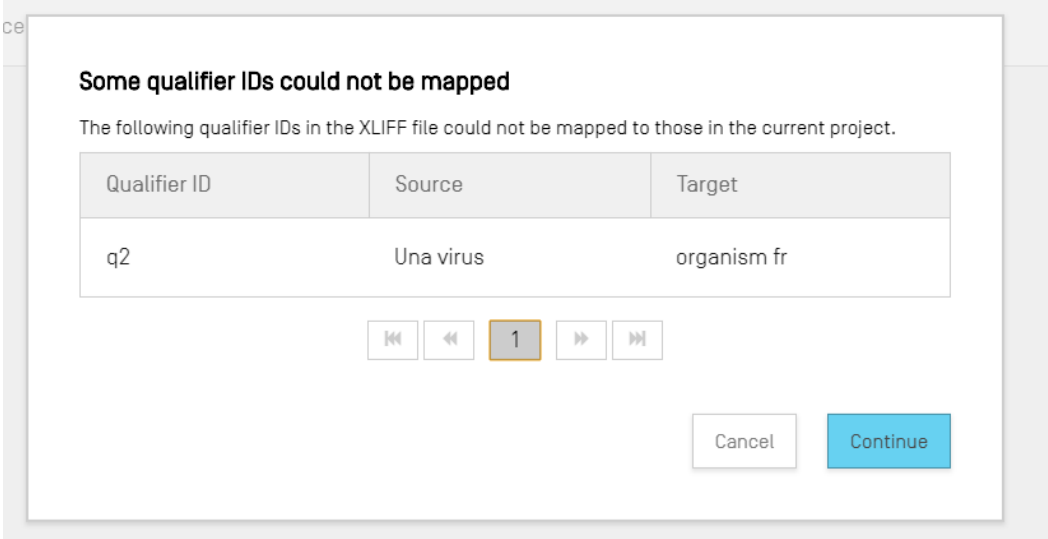


Figure 120 Qualifiers mapping validation

- The system will inform the user of the changes related to the source language and the Qualifiers values:

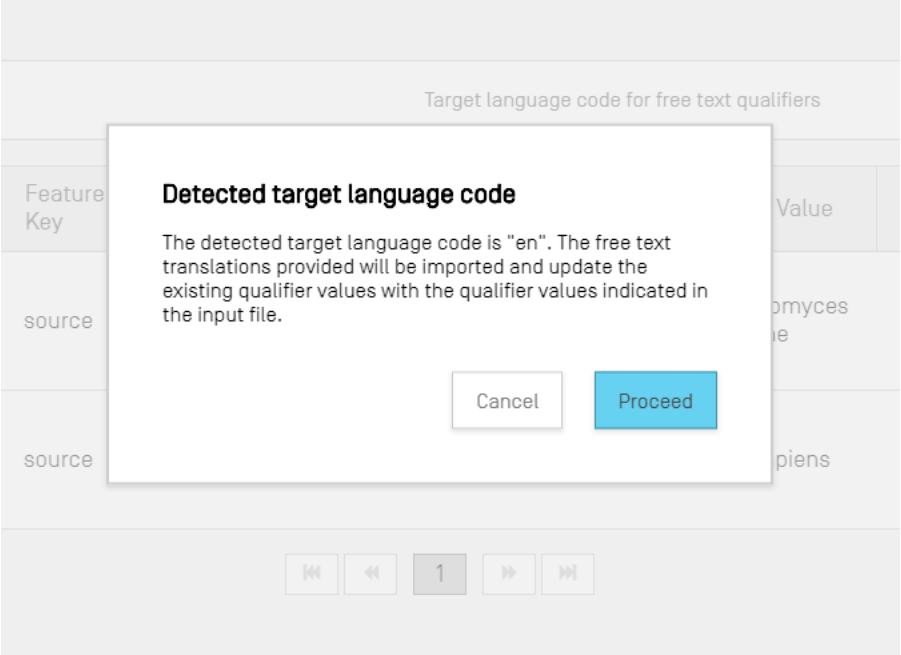


Figure 121 Source language validation

- The system will inform the user of the changes related to the target language and the Qualifiers Translated values:

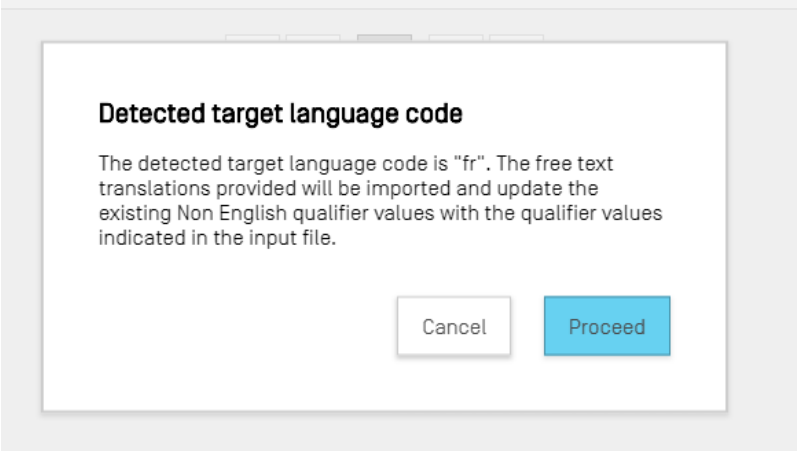


Figure 122 Target language validation

After that we will receive the following banner at the top in blue: 'SUCCESS: THE FREE TEXT QUALIFIER HAS BEEN IMPORTED SUCCESSFULLY' along with an import report displaying in detail the previous and current imported values for the language dependent free text qualifiers.

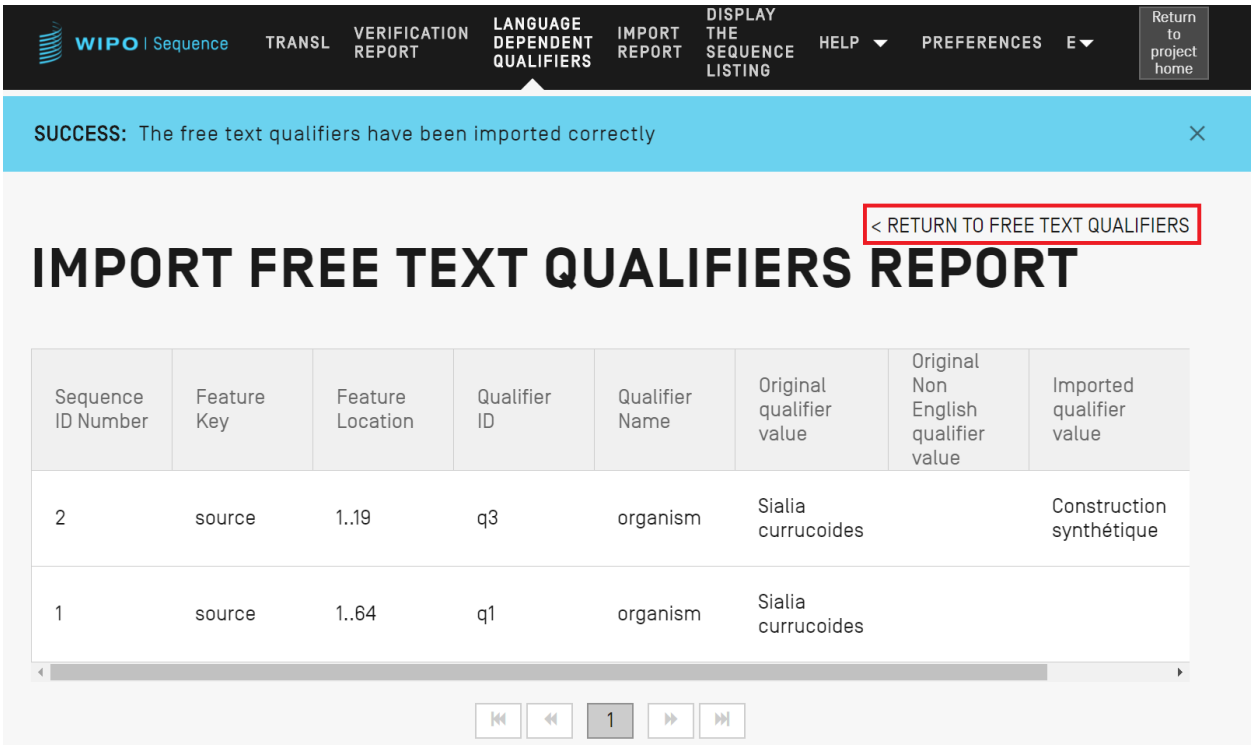


Figure 123 Import Free text qualifier report

User can get back to the Free text Qualifier view by clicking the 'RETURN TO FREE TEXT QUALIFIERS', highlighted in Figure 123.

5.3.2 EXPORT FREE TEXT QUALIFIERS

If the user clicks on the “EXPORT FREE TEXT QUALIFIERS” button at the top of the View, and then in the dialog box, select the file name and location to save the qualifier text file, **all** the free-text qualifiers of the project will be exported and saved to an XLIFF file format.

The file will include:

- The project source language.
- The project target language.
- The free-text qualifiers values.
- The translated qualifier free-text values[10].
- The associated Qualifier and Feature information provided in the Table shown in Figure 123.

This file can be viewed, edited and imported in the tool again after providing the appropriate translation following the steps shown in section **5.3.1**.

5.4 IMPORT REPORT

If a project is imported from a sequence listing (ST.25 or ST.26) or when the user imports multiple sequences from a file (with formats ST.26, ST.25, raw, FASTA or multi-sequence), then the corresponding Import Report will include a Table with all the changes made to the imported data to adapt it to the correct format for inclusion in the project.

If a project was not created by process of importing and no sequence has been imported into the project, the Import Report view will display the banner shown in Figure 124.

[10] The translation of the selected Non English free text language code provided in the project during export

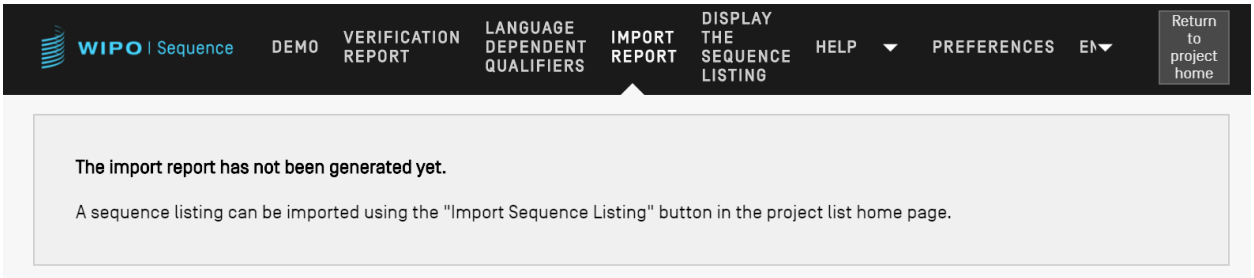


Figure 124 Import report view, no import performed

If an import has been successfully completed and the project has been created, the View could display the following two Tables:

- Import Report Table (see section **5.4.1**)
- Change Data Table (see section **5.4.2**)

5.4.1 Import Report Table

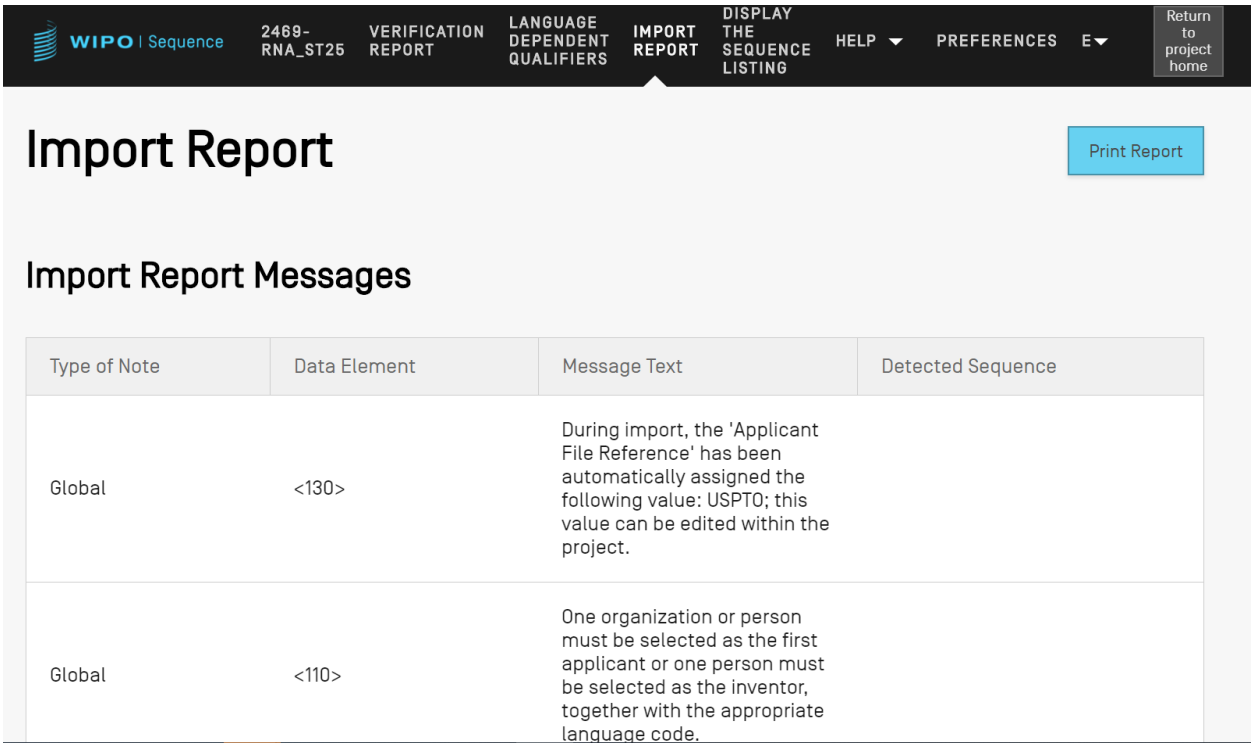


Figure 125 Import report, generated

The Import Report Table is shown only when importing a file results in errors and displays the following columns:

- **Type of note**: “INDIVIDUAL” for a message related to a specific sequence or “GLOBAL” for one or more sequences generally;
- **Data element code**: from the source file, for ST.25 sequence listings;
- **Message text:** Detailed message with information on the identified issue in question and the changes made to rectify it (if any);
- **Detected sequence:** Sequence number of the imported sequence related to the message (when the type is “INDIVIDUAL” otherwise this field is blank).

5.4.2 Change Data Table

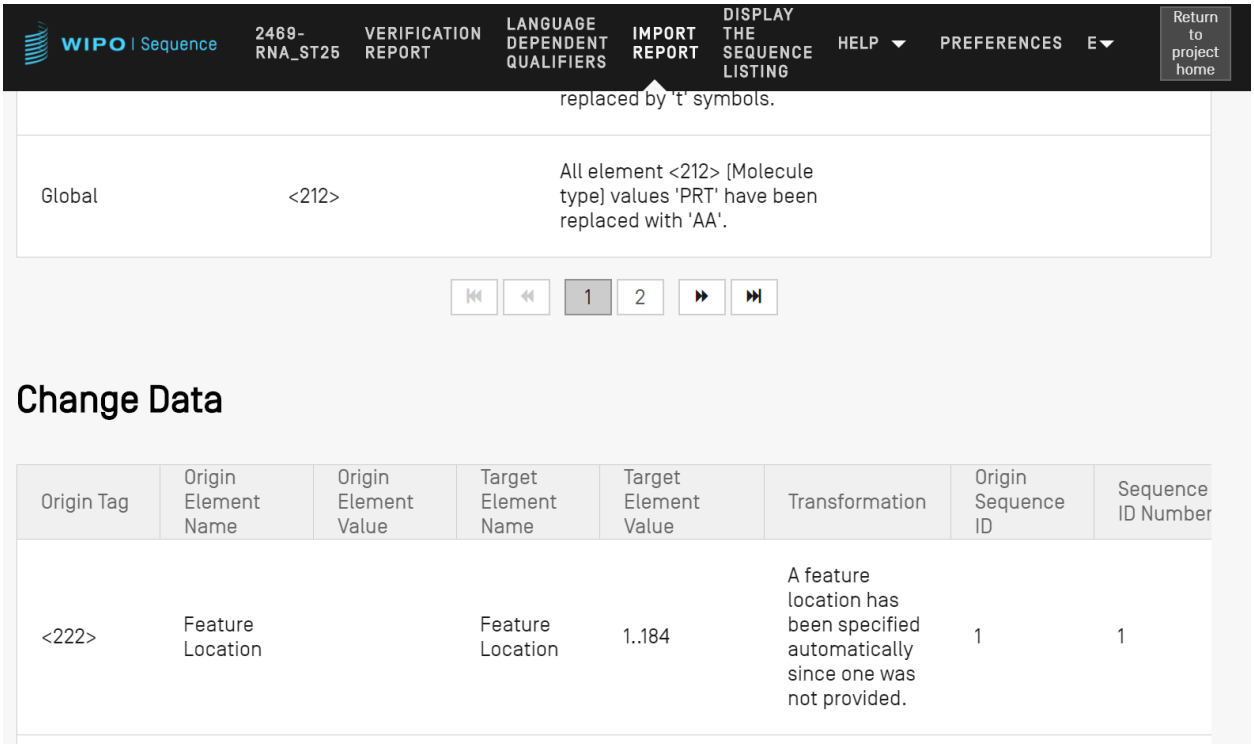


Figure 126 Change Data table

This Table displays any data that has undergone a transformation or change during the importing process. This following data is presented in Table columns (see Figure 126):

- **Origin Tag**: data element code for the element type, for ST.25 sequence listings;
- **Origin Element Name**: corresponding name for the element type;
- **Origin Element Value**: corresponding value of the original element in the source file
- **Target Element Name**: equivalent ST.26 element name where the information is going to be stored in the project;
- **Target Element Value**: value that will be set for the Target Element Name in the project;
- **Transformation**: description of the change(s)/transformation made to the element;
- **Sequence ID Number**: ID number of the relevant sequence of the transformed element in the project.

5.5 DISPLAY SEQUENCE LISTING

WIPO Sequence allows the user to generate a sequence listing in a more human-readable format than XML. When the “DISPLAY THE SEQUENCE LISTING” View is accessed, it will first present a blue banner indicating that the sequence listing file has generated successfully, along with two options to display the aforementioned sequence listing as (.html) or (.txt) format (see Figure 127).

For further details on how to generate a sequence listing, go to section **0**.

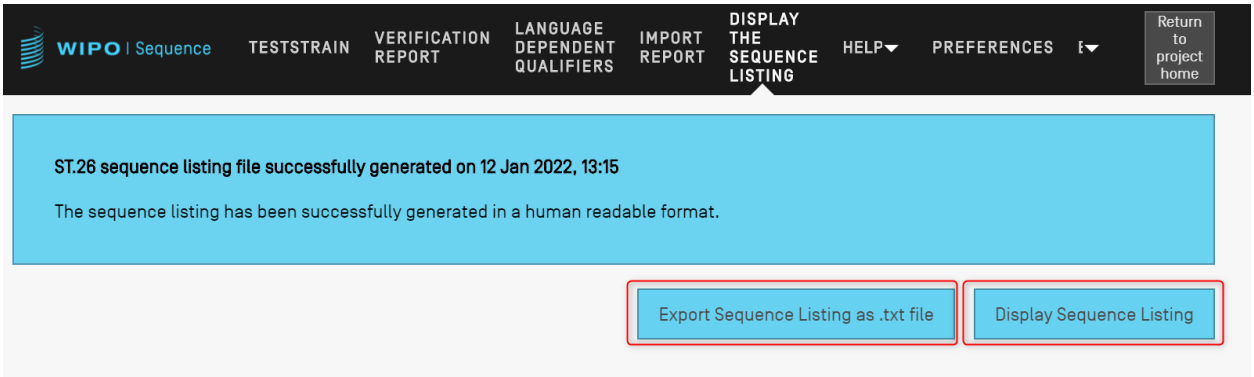


Figure 127 Display Sequence Listing, generated

If a sequence listing has not been successfully generated for a given project, then the Display Sequence Listing View will disable the “Display Sequence Listing” & “Export Sequence Listing as .txt file” buttons and the user will see the error shown in Figure 128.

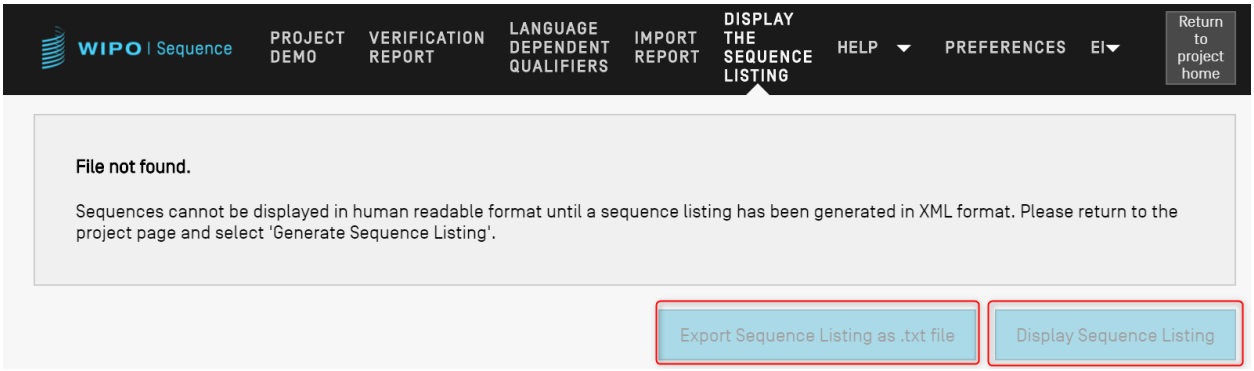
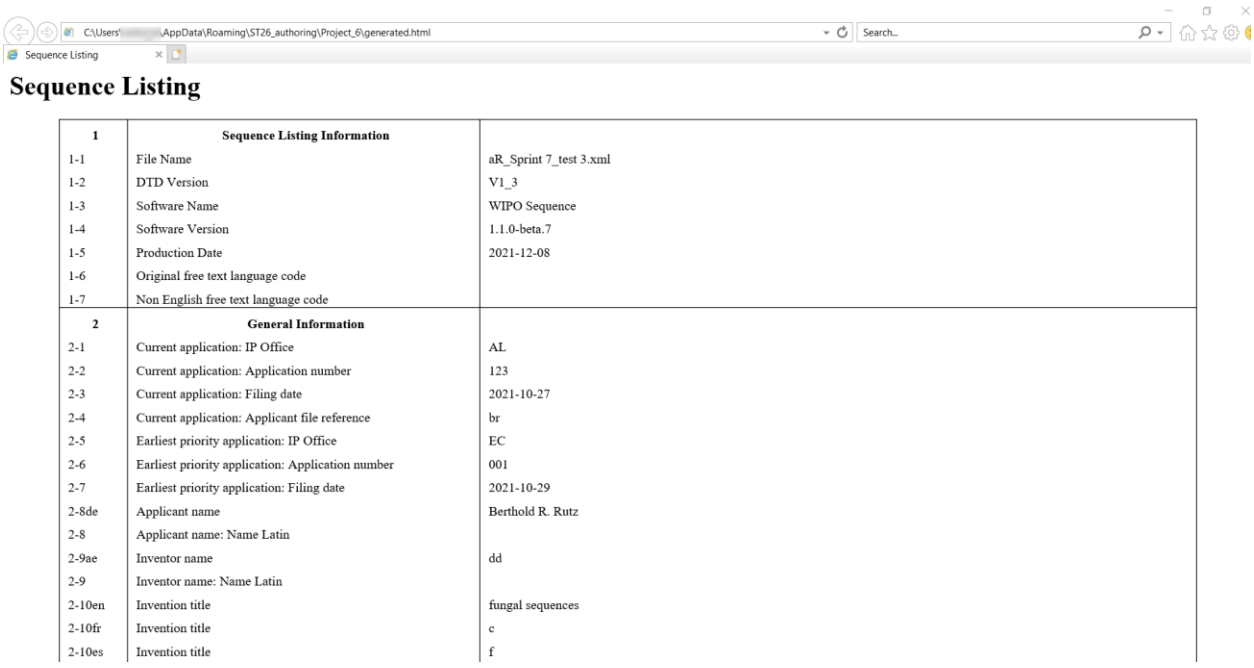


Figure 128 Display Sequence Listing, not generated

When the user clicks on the “Display Sequence Listing” button, a HTML file will be opened in the default browser. This provides a formatted view of the ST.26-compliant XML file so that the values of particular fields are more visible to the user. An example is shown in Figure 129.

Note:

To display the sequence listing in another language, the sequence listing must be generated again. First the new non English free text language code must be indicated in the general information section and then the steps above can be repeated.



Sequence Listing

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

Figure 129 Display Sequence Listing, HTML

Note: the location of the HTML file will be displayed in the navigation bar of the user's browser, in case the user wishes to copy the file into a different location.

When the user clicks on the "Export Sequence Listing as .txt file" button, a txt file will be opened. This provides a formatted view of the ST.26-compliant XML file so that the values of particular fields are more visible to the user. An example is shown in Figure 130.

```
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  ctagctagta  gctgatgatc  atgctagcta  180
        gtagctgatg  atcatgctag  ctagtagctg  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
```

Figure 130 Display Sequence Listing, TXT

If the generated sequence listing, in XML format, is greater than 100Mb in size, instead of displaying the sequence listing in HTML format, the HTML page provided in Figure 131 will be displayed.

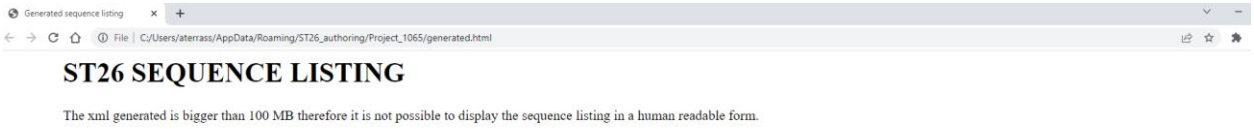


Figure 131 HTML version of sequence listing is too large to display

5.6 HELP

The tool will allow the user to view the help options available to get assistance using this tool (see Figure 132).

Help options direct to information that will:

- Provide a link to the User Manual[11]
- Provide a link to the FAQ[12]
- Provide a link to Contact form for the WIPO Sequence support team
- Provide a link to the WIPO ST.26[13]
- Provide basic information about the WIPO Sequence desktop tool

[11] At the time of publication this points to the English version of the user manual, provided offline

[12] At the time of publication this points to the WIPO ST.26 Implementation FAQ online

[13] At the time of publication this points to the latest version of the Standard online in English

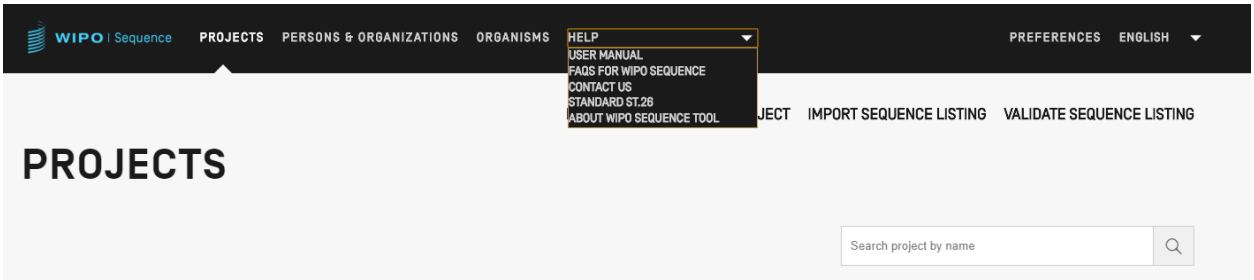


Figure 132 Help

5.7 PREFERENCES

The last View is where the user can set specific properties which apply to all projects which has been detailed previously.

For further details refer to the section **4.4**.

6 FILE FORMATS

6.1 ST.25

For details on the format of WIPO Standard ST.25 files please refer to:

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

6.2 MULTI-SEQUENCE

The multi-sequence format[14] can describe one or multiple sequences, along with their name, the type of molecule and the name of the organism and is one of the allowable formats for import using PatentIn.

The first line of non-blank text is the header and is comprised of the following components:

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

Sequence data begins on the line after the header. A new sequence is delineated by a new line in the file, after the end of the genetic code of the previous sequence. The following is an example of a set of two sequences defined in multi-sequence format.

Example:

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

uuuucuuauuguuucuccuacugcuuaucauaaugauugucguaguggcuuccucaucgucucccccaacc
gccuaccacaacgacugccgcagcggauuacuaauaguaucaccaacagcauaacaaaaagaaugacgaa
gaggguugcugauggugucgccgacggcguagcagaaggaguggcggagggg

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

[14] https://www.uspto.gov/sites/default/files/patents/resources/tools/checker/patentin351_20110214__6_.pdf

```
attgaugtuagtgauguggtautgaugtuagutguagtautgaugtauuaauuaugtggtgagututuga
uatguaautgautugtugutattgaugtuagtgauguggtautgaugtuagutguagtautgaugtauua
auuaugtggtgagututugauatguaautgautugtugutuagt
```

6.3 RAW

This format can only describe one sequence. The genetic code is written in its basic form with no additional information. When imported, molecule type, features and name must be added to the sequence through the tool.

Example:

```
aggatatagatagtatatgatagtatgatatgatgatgtatgtatagtgtagttatga
```

6.4 FASTA

This format contains residues and description and while importing the user has the option to save the description as a note qualifier

6.4.1 FASTA file with one sequence

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

CTCAACAATCTGAAGCATCG
```

See https://www.ncbi.nlm.nih.gov/nuccore/3724029?report=fasta (accessed on 22 May 2017)

[End of Document]