

WIPO Sequence Version 2.3.0 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.

WIPO SEQUENCE Version 2.3.0 USER MANUAL

Contents

1	INTRODUCTION	5
C	OVERVIEW	5
ę	SYSTEM REQUIREMENTS	5
	Installation	6
	Uninstall	7
	Auto update process	8
2	TOOL FUNCTIONALITIES	9
3	TOOL OVERVIEW	13
٦	TOOL MAIN ELEMENTS	13
	View	14
	Sections	14
	Overlay	14
	Tables	15
	Panel Views	
	Edit Panel	16
	Date Picker	17
	PDF Reader	
	Keyboard Navigation	
F	PROJECT HOME VIEW	

WIPO Sequence User Manual	page 3
PROJECT PAGE	
PROJECT DETAIL	
Basic Information	
HELP 24	
PREFERENCES	25
4 TOOL FUNCTIONALITIES	26
PROJECTS HOME	
Create Project	
Import Project	
Import Sequence Listing	
Validate Sequence Listing	
Delete Project	
PERSONS & ORGANIZATIONS	
Create person or organization	
CUSTOM ORGANISMS	
Create Custom Organism	
Export Custom Organisms	
Import Custom Organisms	
SYSTEM PREFERENCES	
PROJECT DETAILS	
Print a project	
Export Project	
Import Information from another Project	
Validate Project	
Generate Sequence Listing	51
General Information	
Sequences	

WIPO Sequence User Manual	page 4
IMPORT REPORT	
Import Report Table	
Changed Data Table	
DISPLAY SEQUENCE LISTING	
5 FILE FORMATS	95
ST.25	
MULTI-SEQUENCE	
RAW	
FASTA	
FASTA FILE WITH ONE SEQUENCE	

1 INTRODUCTION

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 list of the functionality of the tool can be found in section 2 of this document.

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

April 2023

Installation

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 32-bit 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\<username>\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.

Linux

WIPO Sequence is provided as an "AppImage" file (<u>https://appimage.org/</u>) 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.

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.

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).

Uninstall

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

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

OSX

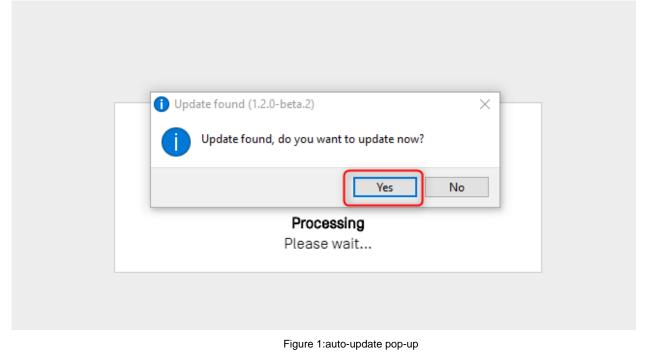
The application can be uninstalled from Finder on OSX, under the Applications section.

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

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.

2 TOOL FUNCTIONALITIES

This section outlines all the functionalities that are implemented by the tool, in the current version of the tool:

Category	Functionality
Projects	Create a project in which data related to one sequence listing is stored
Projects	Generate sequence listing
Projects	Edit the attributes of a project
Projects	Export of free text qualifiers, for the purpose of translation, in XLIFF format
Projects	Export all data stored in a project so that it can be later imported into the same or a different instance of the system (except project metadata)
Projects	Display/export generated sequence listing in human readable format (.html and .txt)
Projects	Import all data stored in a project file (.zip) into a newly created project
Projects	Import data from a ST.25 sequence listing file into a newly created project
Projects	Import data from a ST.26 sequence listing file into a newly created project
Projects	Import data from a FASTA sequence into an existing project
Projects	Import sequence data from a 'multi-sequence' format file
Projects	Import into the current project (target project) the data from another project (origin project)
Projects	Print data from the project

PO Sequence Oser Manua	μ
Projects	Print data from generated ST.26 sequence listing
Projects	Record data that has been changed upon import, in a report indicating the original data and the new changed data
Projects	Verify a ST.26 sequence listing file and list the issues as a verification report containing warning and error messages
Projects	Verify the data stored in a project and list the issues as a verification report containing warning and error messages
Projects	Delete a project
General Information	Add an invention title and its corresponding language code to a project
General Information	Add application information (either current or prior application) to a project
General Information	Add applicant or inventor information to a project
Sequences	Add a source feature and its mandatory qualifiers to a sequence
Sequences	Add feature information to a sequence
Sequences	Add a qualifier to a feature
Sequences	Create a sequence in a project
Sequences	Bulk edit or add features to a range of sequences
Sequences	Bulk skip a series of sequences
Sequences	Bulk delete a range of sequences
Sequences	Create and/or insert a sequence in another position in the listing
Sequences	Import a sequence into an existing project
Sequences	Delete a sequence

WIPO Sequence User Manual

PO Sequence User Manual		page '
Sequences	Automatically create a translation qualifier and optionally its associated amino acid sequence identified by protein_id for a selected CDS feature	
Sequences	Edit qualifier data	
Sequences	Edit sequence data	
Sequences	Edit feature key data	
Sequences	Specify location information for a selected feature	
Sequences	Verify that the residues do not contain invalid symbols	
Sequences	Set the molecule attribute of an amino acid sequence to one of the predefined values	of
Sequences	Translate a nucleic acids sequence according to a specified genetic code table number (by default Genetic Code 1)	
Sequences	Import of free text qualifier source-target pairs in XLIFF forma for the purpose of providing translations in a project	t,
Sequences	Bulk editing of sequence annotation including the qualifier mol_type	
Sequences	Bulk deleting of a range of sequences using bulk edit	
Sequences	Reorder sequences in sequence listing	
Custom Organisms	Add new organism names to the list of organism names store in this system	d
Custom Organisms	Export the list of custom organism names to a text file that can be later on imported into a different instance of WIPO Sequer	
Custom organisms	Import a list of custom organism names from an XLIFF file into an existing project	0
Person/Organizatio n	Store in the system information about an applicant or inventor (e.g., name, its corresponding language code and its translation or transliteration into Latin characters (if applicable), address	

April 2023

WIPO Sequence User Manual

	etc.) so that it can be used later in various projects	
Person/Organizatio n	Create a new person or organization name	
System Preference	Adjust certain preferences of the tool which effect all projects generated by that instance	

3 TOOL OVERVIEW

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.

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
 - o 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
 - o Display Sequence Listing View

WIPO Sequence PI	ROJECTS PERSONS & ORGANIZATION	S ORGANISMS HELP	*		PREFERENCES ENGLIS	•
		NEW PROJECT IMP	ORT PROJECT IMPORT SEQ	UENCE LISTING	VALIDATE SEQUENCE LIS	TING
PROJECT	S					
			Search	project by name		Q
Project name :	Applicant file reference :	Applicant name :	Invention title -	Status :	Creation Date :	
Test preferences	applicant	PETIT, Robert	invention	modified	2021-12-09	
DEMO	br	Berthold R. Rutz	fungal sequences	modified	2021-08-12	
ST26T-2540	30610/44556C1-R2	BioMarin Pharmaceutical Inc.	Variants of C-Type Natriuretic Peptide	modified	2021-09-28	

Figure 2: Project Home View

WIPO Sequence User Manual

View

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

Sections

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

	WIPO Sequence	PROJECT_TITLE	VERIFICATION REPORT	LANGUAGE DEPENDENT QUALIFIERS	IMPORT REPORT	DISPLAY THE SEQUENCE LISTING	HE▼	PREFERENCES	•	Return to project home
GE	NERAL INFORMATION	SEQUENCES								
	Original free text langua Automatically add a trai created Off		en a CDS feature i			xt language ci	ode			Ø
<mark>2</mark> ▶ (ENERAL IN	FORMATI	DN							
⊳ S	EQUENCES									
										3 ↑

Figure 3: Project Details View

Within the Project Detail View, shown in Figure 3, there are two 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).

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 4.

O Sequence User Manual				page 15
GENERAL INFORMATION SEQUENCES	PRIORITY IDENTIFICATION			
PRIORITY IDENTIFICATION	Add Excitized Printity Application			
Add Hadlest Prodty Application	IP Office*	•	Filing dete yypyy-MM-riti	=
APPLICANT & INVENTOR Add Applicant Add Inventor	Application monther*		Selected Entliest Printly Application 1	
INVENTION TITLE			Carroel	Add Earliest Provity Application
Ade Invention title	APPLICANT & INVENTOR			
	Add Appliand Add Intentor			
	INVENTION TITLE			

Figure 4: Overlay

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

Tables

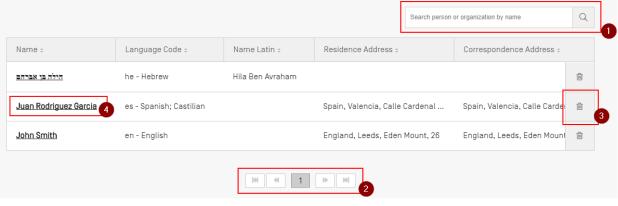


Figure 5: Example Table

An example of a Table is shown in Figure 5. 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).

Name :		Name Latir		Residence Address o	Correspondence Address :	
הילה בן אברהם	he - Hebrew	Hila Ben Av	/raham			
Juan Rodriguez Garcia	es - Spanish; Castilian			Spain, Valencia, Calle Cardena	Spain, Valencia, Calle Carde	
John Smith	en - English			England, Leeds, Eden Mount, 26	England, Leeds, Eden Moun	
			1			
Name* Juan Rodriguez Garcia				Residence Address Spain, Valencia, Calle Cardenal Benlloc	.h, 10	
Language* es - Spanish; Castilian			*	Correspondence Address Spain, Valencia, Calle Cardenal Benlloc	.h, 10	
Name Latin						

Figure 6: Editing a table

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

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.

GENERAL INFORMATION



Application Identified Before the assignment of the application IP Office AI - Anguilla number Applicant file reference 4342

Application number 32424 Filing date 2022-01-03

Figure 7: Example Panel View

As shown in Figure 7, when the button with the pencil icon (1) is clicked on a panel view, the Panel View is replaced with an Edit Panel.

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

0

8).

Note:

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

PPLICATION IDENTIFICATION		
IP Office GB		-
Application Identification*	Applicant file reference* ABC1234 Application File Reference is a mandatory field when 'Before the as number' is provided.	ssignment of the application
Application number 98968268463829	Filing date 2019-02-21	í.

Figure 8: Editing a Panel

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 9.

	Yea 20	ar)19	*		Month Feb	n ,	Ŧ
	27	28	29	30	31	01	02
	03	04	05	06	07	08	09
	10	11	12	13	14	15	16
	17	18	19	20	21	22	23
Approved Neuropeancer	24	25	<u>26</u>	27	28	01	02
A80224	03	04	05	06	07	08	09
Approximation and an exercise here and approximation of the provider.	<]		Today	/		>
Filing date 2019-02-22							



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 ("YYY-*MM-DD"*) (3).

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 thefile and save it, the user must click on the download icon at the top right of the viewer (1), shownin Figure 10.



Figure 10: PDF Viewer

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 in Figure 11:

FEATURES		
Feature Key	Location	Qualifiers
source	164	mol_type = genomic DNA organism = Sialia currucoides
	K K 1	₩ ₩
		an Ealth facture butter

Figure 11: Focus on Edit feature button

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

WIPO Sequence User Manual

For example, when creating a new project, there is only one mandatory field: Name. The cursor and focus are already set toproject name field as shown in Figure 12 below:

WIPO Sequence	NEW PROJECT H	IELP	•	PREFERENCES	ENGLISH	•	Return to project home
Name*			Description				
					Can	cel	Save

Figure 12: Cursor focus on the mandatory field

PROJECT HOME VIEW

The project home View of the tool consists of 3 main Views:

The *Project* View (see Figure 13), the *Persons & Organizations* View (see Figure 14) and the

Organisms View (see Figure 15). There is also the Preferences (see Figure 19 which apply to all projects in the top right-hand corner.

PROJECTS			NEW PROJECT IMPORT PROJECT	IMPORT SEQUENCE	LISTING VALIDATE SEQUENCE LISTIN
				Search project b	y name Q
Project name :	Applicant file reference :	Applicant name :	Invention title :	Status :	Creation Date :
<u>cds feature</u>	15123-W0-PCT[2]	Novozymes A/S	Polypeptides Having Beta- Glucanase Activity and Polynucleotides Encoding Same	modified	2021-09-07
<u>160K 500 SEQ</u>	A400: 66076	University of Tokyo	COPOLYMER INCLUDING UNCHARGED HYDROPHILIC BLOCK AND CATIONIC POLYAMINO ACID BLOCK HAVING HYDROPHOBIC GROUP IN PART OF SIDE CHAINS, AND USE THEREOF	modified	2021-09-07
160K 100 SEQS	A400: 66076	University of Tokyo	aaaaaaaa	modified	2021-09-06
<u>cds feature</u>	15123-W0-PCT[2]	Novozymes A/S	Polypeptides Having Beta- Glucanase Activity and Polynucleotides Encoding Same	invalid	2021-09-02
<u>cdsFeatures</u>	15123-W0-PCT[2]	Novazymes A/S	Polypeptides Having Beta- Glucanase Activity and Polynucleotides Encoding Same	modified	2021-09-01
		Figure 13: Project	te Homo View		

WIPO Sequence P	PROJECTS PERSONS & ORGANIZATIONS ORGA	NISMS HELP	+	PREFERENCES	INGLISH
				CREATE NEW PERSON OR ORG	ANIZATIO
PERSONS	& ORGANIZATI	DNS			
				Search person or organization by name	Q
Name :	Language Code =	Name Latin o	Residence Address :	Correspondence Address :	
<u>Джо, Смит</u>					
<u>Джо. Смит</u> Д <u>жейн, Эйр</u>					

Figure 14: Persons & Organizations View

			page 21
WIPO Sequence PROJECTS PERSONS & ORGANIZATIONS ORGANISMS HELP	~	PREFE	ERENCES ENGLISH 🔻
	ORT CUSTOM ORGANISMS	IMPORT CUSTOM ORGANISMS	CREATE NEW ORGANISM
ORGANISMS			
		Search organism by name	Q
Name :	Description		
test organism			
<u>Demo Organism</u>			
B	bbb		
Δ	aaa		
Jec			

Figure 15: Organisms View

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 16:

- 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, WIPO Standard ST.26 and WIPO Sequence and ST.26 Knowledge Base (6)
- 7. Preferences View (7).

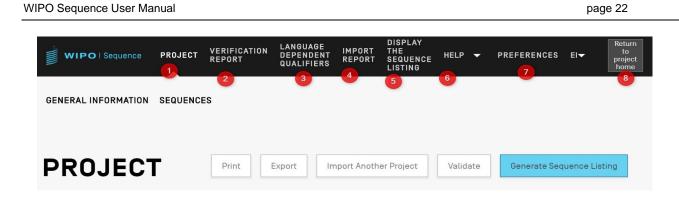


Figure 16: Project Details toolbar

To return to the Projects home page, the user can click on the "Return to projecthome" button (8) at the right end of the header/toolbar.

PROJECT DETAIL

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 17.

This section contains:

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

PO Sequence Use	er Manua	al						p	age 24	
WIPO Sequence	ST26T- 2136	VERIFICATION REPORT	LANGUAGE Dependent Qualifiers	IMPORT Report	DISPLAY THE SEQUENCE LISTING	HELP	•	PREFERENCES	ENGLI	Return to project home
GENERAL INFORMATION	SEQUEN	CES								
ST26T-2	136			Print	Export Import Ar	nother Project	Validate	Generate	e Sequence L	isting
					Creation date 2023-04-06					

Figure 17: Basic information section

HELP

The tool will allow the user to view the help options available to get assistance using this tool (see Figure 18).

Help options direct to information that will:

- Provide a link to the User Manual¹
- Provide a link to ST.26 Knowledge Base ²
- Provide a link to Contact form for the WIPO Sequence support team.
- Provide a link to the WIPO ST.26³
- Provide basic information about the WIPO Sequence desktop tool.

¹ At the time of publication this points to the English version of the user manual, provided offline or in remaining 9 PCT languages online

² At the time of publication this points to the WIPO Sequence and ST.26 knowledge base

³ At the time of publication this points to the latest version of the Standard online in English, French or Spanish

Note:

Packaged with the latest version of WIPO Sequence is the user manual and WIPO ST.26 in English only. When a user selects a different language for the GUI, they will need to be connected to the internet in order to access other language versions of both the Standard and the user manual. However, the link provided will be to the language version which matches their selection.

WIPO Sequence	PROJECTS	PERSONS & ORGA	NIZATIONS	ORGANISMS	HELP USER MANUAL WIPO SEQUENCE KNOWLI	EDGE BASE	•		PREFERENCE	S EN(GLISH	•
				NEW PRO	CONTACT US STANDARD ST.26 DJ ABOUT WIPO SEQUENCE		SE	EQUENCE LISTING	VALIDATE SEC	UENCE	ELISTIN	١G
PROJEC	TS											
							Sear	rch project by name			Q	
Project name o	Applicant f	ile reference o	Applicant	name 🗧	Invention title 0	Status o		Creation Date o	Last r	nodifie	d -	
			Fig	gure 18: H	lelp dropdown mer	าน						

PREFERENCES

The last View is where the user can set specific properties that apply to all projects (see Figure 19). Specific details on what each of these parameters are can be found below in the relevant 'Tool functionalities' section.



4 TOOL FUNCTIONALITIES

PROJECTS HOME

This section details the different functionalities 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.

Create Project

To create a new project, the user must begin from the main Projects Home View (see Figure 20).

WIPO Sequence	PROJECTS	PERSONS & ORGANIZATIONS	ORGANISMS	HELP	.				PREFERENCES	ENGLISH	•
						NEW PROJECT IMPORT PROJECT	IMPOR	RT SEQUENCE LISTING	VALIDATE SEQU	ENCE LISTI	ING
PROJEC [®]	TS										
								Search project by name		C	2
Project name :		Applicant file reference :		Applicant nam	e :	Invention title :	Sta	tus o	Creation Date :		

Figure 20: Create new project Step 1

1) Click on the "NEW PROJECT" link at the top of the View indicated in Figure 20. In the following view, the tool will request a 'Name' (mandatory) and a 'Description' (optional).

WIPO Sequence NEW PROJECT HELP	•	PREI	FERENCES	ENGLISH	•	Return to project home
Name*		Description				
				Cance	əl	Save

April 2023

Figure 21: Create new project Step 2

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 22.

	WIPO Sequence PROJECTS	PERSONS & ORGANIZATIONS ORGANISMS HEL	p 🗸			PREFERENCES ENG	olish 👻
F	PROJECTS			NEW PROJECT IMPORT PROJECT	IMPORT SEQUENCE LISTIN	IG VALIDATE SEQUENCE	LISTING
					Search project by name		Q
	Project name o	Applicant file reference o	Applicant name o	Invention title :	Status :	Creation Date o	
	<u>cds feature</u>	15123-W0-PCT[2]	Novozymes A/S	Polypeptides Having Beta- Glucanase Activity and Polynucleotides Encoding Same	modified	2021-09-07	
	160K 500 SEQ	A400: 66076	University of Tokyo	COPOLYMER INCLUDING UNCHARGED HYDROPHILIC BLOCK AND CATIONIC POLYAMINO ACID BLOCK HAVING HYDROPHOBIC GROUP IN PART OF SIDE CHAINS, AND USE THEREOF	modified	2021-09-07	

Figure 22: New project shown in summary

Import Project

This functionality allows the import into the tool of a previously exported project as shown in Figure 53. To import a project file, the user must begin from the Projects Home View as shown below in Figure 23.

WIPO Sequence P	PROJECTS	PERSONS & ORGANIZATIONS	ORGANISMS	HELP	•				PREFERENCES	ENGLISH	•
PROJECT	S					NEW PROJECT	ІМРО	NRT SEQUENCE LISTING	VALIDATE SEQUE	ENCE LIST	TING
								Search project by name		(Q
Project name :		Applicant file reference o		Applicant n	name o	Invention title :	Sta	atus :	Creation Date o		
cds feature		15123-W0-PCT[2]		Novozymes	s A/S	Polypeptides Having Beta- Glucanase Activity and Polynucleotides Encoding Same	m	odified	2021-09-07		

Figure 23: Import Project Step 1

1) Click on the "IMPORT PROJECT" link at the top of the view as indicated in Figure 23.

WIPO Sequence User Manual

					_	
		NEW PR	OJECT IMPORT PROJECT	IMPORT SEQUENCE LIS	ring valida [.]	TE SEQUENCE LISTING
					1	Upload file [.zip]
	Open			×		
	\leftrightarrow \rightarrow \land \uparrow \square \rightarrow This PC \Rightarrow D	cuments > WIPO_ST26	ڻ ~		Cancel	
	Organise 🔻 New folder			III 🕶 🔟 🚱	-	
PROJEC	🖈 Quick access	Name	Date modified	Туре	liz	
	everis	2 🗱 tets_1617812352954.zip	07/04/2021 17:19	WinRAR ZIP archive		
	OneDrive - everis					
	This PC				e	Q
	Network					
Project name :					Cre	
0					201	21-05-12
*					20.	
		٢			>	
	File name: tets	617812352954.zip	~	ZIP (*.zip) V		

Figure 24: Import project dialog box

- 2) On the Overlay screen shown in Figure 24, click the "Upload file [.zip]" button (1).
- In the dialog box that opens, shown in Figure 24, select the project file to be imported (2 & 3).

WIPO Sequence PROJECTS PERSONS & ORGANIZATIONS ORGANISMS HELP				
NEW PROJEC	IMPORT PROJECT	IMPORT SEQUENCE LISTING	VALIDATE SEQUEN	CE LISTING
st28t-2576 (1].xmL1840106895750.zip			Upload file	e [.zip]
■ Enter the sequences to be imported				
Sequence Number [ID]	Position			
1				
2				
H4 44 1 H> H4				
Select Range of Sequence IDs. 1-2				
Use commas to separate individual sequences (for example: "2,5,8"); use a dash to indicate a range of sequences (for example: "2-8").				
			Cancel Impo	rt Project

Figure 25: Import project, select sequence range

4) If "Select Range of Sequence IDs" 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 of Sequence IDs" checkbox (1) and enter the ID numbers of the desired sequences in the appropriate field (2) shown in Figure 25. A single sequence can be entered, as well

April 2023

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 25.

If the project is successfully imported, the following blue banner and message will appear at the top of the View as shown in Figure 26.

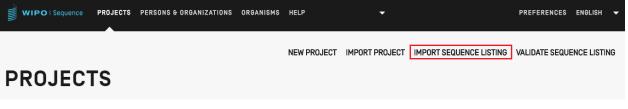
WIPO Sequence PROJECTS PERSONS & ORGANIZATIONS ORGANISMS HELP	PREFERENCES ENGLISH 🗸
SUCCESS: The project has been imported successfully	×
Figure 26: Successfully imported a project, blue bapper	

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.

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 27.





 In the Overlay that opens, shown in Figure 28,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).

WIPO Sequence User Manual

Organice New folder IEF IFF		$\leftarrow \rightarrow \cdot \uparrow \square \rightarrow$ This F	PC > Document	WIDO ST26	✓ ♂ Search WIPO	5726		
validsts.xml Documents Name Date modified Type Project Name * Downloads Dissicxml 02/03/2021 15:37 XML Project Name * Pitures SIZE1-T96-validxml 02/03/2021 15:39 XML Valid test Desktop Desktop Desktop Desktop Sample Files WPO Desktop Desktop Desktop WIPO_SIZE Size Desktop Desktop Desktop Desktop Desktop Desktop Desktop Desktop Select the general information corr WPO_SIZE Desktop Desktop This PC Network Cancel Import Sequence Pitenme Desktom Desktom Desktop This PC Network Cancel Import Sequence			/C > Document	; > WIPO_S126			ISTING VALIDATE SEQUE	ENCE LISTING
validsts.xml 2009/2021 15:37 VML Project Name* valid test 2009/2021 15:37 VML Str257-1796-validxml 2009/2021 15:39 VML Str257-1796-validxml 2009/2021 15:37 Test Str257-1796-validxml 2009/2021 15:39 VML Str257-1796-validxml 2009/2021 15:39 VM		-			855	• 🔳 🕜		
work work <th>validsts.xml</th> <th></th> <th>* ^</th> <th>Name</th> <th>Date modified</th> <th>Тури</th> <th>Upload ST.25 Ltxt1 or ST.26</th> <th>(.xml) file</th>	validsts.xml		* ^	Name	Date modified	Тури	Upload ST.25 Ltxt1 or ST.26	(.xml) file
Project Name.* Music bttl.7796 tthuman 1.tt 02/03/2021 15:37 Text Valid test Deskop bttl.7796 tthuman 2.txt 02/03/2021 15:37 Text Select Range of Sequence IDs WIPO_ST26 intervision intervision intervision Select the general information on intervision intervision intervision intervision Image: Select the general information on intervision intervision intervision intervision Image: Select the general information on intervision intervision intervision intervision Image: Select the general information on intervision intervision intervision intervision intervision Image: Select the general information on intervision intervision intervision intervision intervision Image: Select the general information on intervision intervision intervision intervision intervision Image: Select the general information on intervision intervision intervision intervision intervision Image: Select the general information on intervision intervision intervision intervisi	- Charles and a charles and a charles a charle	and the state of the	1	🗐 basic.xml 🦷	02/03/2021 15:3	7 XMI		interior into
valid test Deskop Select Range of Sequence IDs Select the general information construction Select the general information construction File name basic.xml File name basic.xml Select Xamp of Sequence IDs Select Xa		and the second se	1					
Select Range of Sequence IDs Sample Files WIPO_St26 Belect the general information on This PC Network File name basic.xml Select xml("tst;"xml)		and the second se	*					
Select Range of Sequence IDs WPO WPO_5T26 evris OneDrive - evris oneDrive - evris This PC Network File name basic.xml File name basic.xml	valid test	A CONTRACTOR DE CONTRACTOR		txt_1796 txt human 2.txt	02/03/2021 15:3	.9 Text	¢.	
Select the general information on Select the general information on This PC Network File name basic.xml tx xml("1st;"xml)		a service concernation						
Select the general information con Select the general information con This PC This PC File name basic.xml Let xml("dst",xml) Cancel Import Sequence	Select Range of Sequence IDs	1						
Select the general information con Image: Conceptive - events Image: This PC Image: Conceptive - events Image: File name (basic:xml) Image: Conceptive - events		WIPO_ST26						
Cancel Import Sequence File name basic.xml Ltt xml("tst;"xml)		i everis						
This PC This PC The name basic xml The name	Select the general information cor							
import Sequence File name basic.xml v intra xml("sixt"xml)								
File name [basic.xm] v [.txt xml("txt"xml) v		This PC						_
		Network	~	٢		>	Cancel Import Sequ	ence Listing
PROJECTS		File nam	ne: basic.xml		v .txt .xml (*.txt;*.xn	nl) ~		
	DOILOTO				Open	Cancel		
	RUJEUIS				open	curres		

Figure 28: Import sequence listing, dialog box

Two primary checkboxes shown in Figure 29 will allow the user to specify which sections to be imported into the new project, "Select Range of Sequence IDs" and "Select the general information contents to be imported".

	NEW PROJECT	IMPORT PROJECT	IMPORT SEQUENCE	LISTING	VALIDATE SEQUENCE LISTING
st261-2624.xml				Upload S	ST.25 [.txt] or ST.26 [.xml] file
Project Name *					
Sequence Number [ID]		Position			
[1]		1			
M 4 1 >> M	И				
Select Range of Sequence IDs.					
Use commas to separate individual sequences (for example: "2,5,8"); use a dash to indicate a range of sequences (for example: "2-8").				
Select the general information contents to be imported					
				Cancel	Import Sequence Listing

Figure 29: 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.

April 2023

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 28 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.

Project N Insectid	lame Ial Proteins		
Selection	it Range Sequences		
✓ Select	t the general information contents to be imported		
	Element	Origin Element Value	Target Element Value
•	Application Identification		IP Office = GB Application number = 34892756 Filing date = 2019-05-02
•	Applicant File Reference		Applicant file reference = ABCD1234567
	Earliest Priority Application Identification		IP Office = GB Application number = 128432643875345 Filing date = 2019-05-01
	Applicant Name		Name = James Wilson
	Invention Title Bag		Invention title = Insecticide protein, Language code = en
The da	ta for the selected attributes will be overwritten		
			Cancel import Project

Figure 30: 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 30.

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

If the Sequence listing was imported correctly, the Changed Data Table shown in Figure 31 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.

nport	E Report						Print Repo
Origin Tag	Origin Element Name	Origin Element Value	Target Element Name	Target Element Value	Transformation	Origin Sequence ID	Sequence ID Number
	Sequence	u	Sequence Residue		All 'u' symbols in sequence are not permitted and should be changed to 't' symbols and that either a modified, base (uracil in DNA) or misc_feature (combined RNA/DNA molecule) is needed.		10

Figure 31: Import sequence listing, Changed Data

If the file format was ST.25, then the Import Report View will include an Import Report Table first, as well as the Changed Data Table. An example of the Changed Data report is shown in Figure 31 and an example Import Report is shown in Figure 32.

WIPO Sequence	PROJECT Demo	VERIFICATION REPORT	LANGUAGE DEPENDENT QUALIFIERS	IMPORT REPORT	DISPLAY THE SEQUENCE LISTING	HELP	*	PREFERENCES	ENGLIS V	Return to project home	2
Import Ro	eport								1 Prir	nt Report	
Import Repo	t Messa	IGES Data Element		Message Te	xt		Detected	Sequence			
Global		<400>		be changed should be us DNA sequen with a 'note'	'u' is not permitted an to a 't'. A 'modified_ba sed to identify a uracil ce. A 'misc_feature' al ' qualifier should be us h fragment of a DNA/R ence	ise' in a ong ed to					
Global				qualifier 'mo	ide appropriate value(s ol_type' of the following s prescribed by Annex	SEQ ID					

Figure 32: Import sequnce 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 32.

For instructions on how to download the PDF file, see Figure 10.

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 33.

WIPO Sequence PR	DJECTS PERSONS & ORGANIZATIONS	ORGANISMS HELP	•		PREFERENCES ENGLIS	н 🔻
ERROR: An error occurred	while importing the ST.25 sequenc	e listing.				×
PROJECTS	5	NEW PROJECT	IMPORT PROJECT IMPO	RT SEQUENCE LISTING	VALIDATE SEQUENCE LIS	
Project name o	Applicant file reference	Applicant name	Invention title	Search project by name Status	Creation Date :	Q
<u>gv 42 xgv 42</u>	AB123	Tom Jons	Copolymer including uncharged hydropfilic block	new	2021-07-07	

Figure 33: Import ST.25 sequence listing, red banner

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 sequences 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.

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 34.

WIPO Sequence PROJECTS	PERSONS & ORGANIZATIONS	ORGANISMS HE	LP 🗸				PREFERENCES ENG	GLISH 🔻
				NEW PROJ	ECT IMPORT PROJECT	IMPORT SEQUENCE LISTIN	IG VALIDATE SEQUENCE	LISTING
PROJECTS								
						Search project by name		Q
Project name o	Applicant file reference :		Applicant name o	Inventior	title :	Status :	Creation Date :	

Figure 34: Validate sequence listing

Next, the user must click on the "Upload file ST.26 [.xml]" button (1) shown in Figure 35, and then select the file in the dialog box (2). Lastly, click the "Validate Sequence Listing" button (3).

Sequence	e User Manual				page 34
in Mil	PO Sequence PR	OJECTS PERSONS & ORGANIZATIONS	ORGANISMS HELP		RENCES ENGLISH
		NEW PROJECT IMPORT PI	ROJECT IMPORT SEQUENCE L		SEQUENCE LISTIN
					d file ST.26 [.xml]
	Open		×	3	
	← → • ↑ 📜 « Do	ocu > sequences JDD 🗸 🗸	Ca Search sequences JDD	ancel Validat	e Sequence Listing
DD	Organise • New folde	er	III - 🔳 🕐		
PR	📕 reports 🛛 🖈 🔨	Name	Date modified		
	2652	2021097599 (2)	9/9/2021 3:00 PM		
	sequences JDD	16515000_1_1 (4)	9/9/2021 3:00 PM		
	WIP		9/9/2021 3:00 PM		
	KFQ_11	txt_st25_artificial_unknowntxt	9/9/2021 3:00 PM		
		adt 🦪	9/9/2021 3:00 PM ect by	name	Q
	📓 NTT DATA EMEAL	AEPAI_FD_3_XFD_4 (1)	9/9/2021 2:14 PM		
	合 OneDrive - NTT D,	a) basic	9/9/2021 1:51 PM		Creation
Proje	📜 Microsoft Teams 🗸	ST_26_long_amended DTD version	9/9/2021 1:39 PM 🗸 itle 🗧	Status o	Date
<u>2404</u>	File na	me: basic	nl v Open Cancel	modified	2021-11-08

Figure 35: Validate sequence listing, dialog box

If the sequence listing passes validation, a banner will appear as shown in Figure 36:

SUCCESS: The ST.26 project has been verified successfully	×

Figure 36: Banner indicating validation completed successfully

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 37.

page 35

Verification report

Verification Report Information

	Production Date	2022-02-12
Ľ	Verification report (XML)	C:\Users\aterrass\AppData\Roaming\ST26_authoring\QT05.xml
	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: xi,o,ri,e,x,y.a. Only printable characters (including the space character) from the Unicode Basic Latin code table (except the reserved characters) are permitted.	молекула	1
ERROR	R 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 37: Validate sequence listing, verification report

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 (see Figure 37).

Note: 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.dtd⁴.

Delete Project

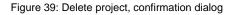
To delete a project, the user must begin from the Projects (home) View, shown in Figure 38.

PROJECTS			NEW PROJECT IMPORT PROJECT	IMPORT SEQUENCE LISTIN	G VALIDATE SEQUENCE LISTING
				Search project by name	٩
Project name o	Applicant file reference o	Applicant name o	Invention title o	Status o	Creation Date :
<u>st26t-2624</u>	es123	ALVIZO, OSCAR Nuñez	ST26T- 2624 ES	new	2022-01-12
<u>cds feature</u>	15123-W0-PCT[2]	Novozymes A/S	Polypeptides Having Beta- Glucanase Activity and Polynucleotides Encoding Same	modified	2021-09-07

Figure 38: Delete project, project home view

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

PROJECTS			NEW PROJECT	MPORT PROJECT	IMPORT SEQUENCE LIS		
	Applicant file reference o	Applicant name o					
<u>st26t-2624</u>	es123				new	2022-01-12	
<u>cds feature</u>	15123-W0-PCT[2]	Delete Project Are you sure that you want to permanently project?	delete this y a	ng Beta- ⁄ and ncoding	modified	2021-09-07	
<u>160K 500 SEQ</u>	A400: 66076	University of Tokyo	BLOCK AND CATI POLYAMINO ACID HYDROPHOBIC 6 OF SIDE CHAINS, THEREOF	BLOCK HAVING ROUP IN PART	modified	2021-09-07	



⁴ This is the current version of the DTD at the time of publication of this document

In the pop-up shown in Figure 39, click "Delete" to confirm that you want to delete the selected project.

PERSONS & ORGANIZATIONS

This section details the functionalities provided in the Person & Organizations View.

Create person or organization

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

				CREATE NEW PERSON OR OR	GANIZATION
ERSON	IS & ORGANIZ	ATIONS			
				Search person or organization by name	Q
			D : I A I I	Correspondence Address :	
Name o	Language Code o	Name Latin :	Residence Address o	our opportabilition radiced v	

Figure 40: Create new person or organization Step 1

First, the user must click on the "CREATE NEW PERSON OR ORGANIZATION" link at the top of the View, as shown in Figure 40.

WIPO Sequence CREATE NEW PERSON OR ORGANIZATION HELP		•	PREFERENCES	ENGLISH	•	Close
Name*		Residence Address				
Language*	v	Correspondence Address				
Name Latin						
				Can	cel	Save

Figure 41: Create new person/organization Step 2

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.

CUSTOM ORGANISMS

To create, edit, import, export or delete Custom Organisms, the user must begin from the Organisms View. Details on how to create, export or import custom organisms are provided below in Figure 42.

Create Custom Organism

WIPO Sequence	PROJECTS	PERSONS & ORGANIZATIONS	ORGANISMS	HELP	•	PREFERENCES ENGLISH 🔻			
ORGANIS	SMS				EXPORT CUSTOM ORGANISMS	IMPORT CUSTOM ORGANISMS	CREATE NEW ORGANIS	SM	
						Search organism by name	Q		
Name :				E	Description				
test organism				C	Description example				
Demo Organism				[Description Demo example				
B				b	bb				

Figure 42: Create new organism Step 1

To create a new custom organism, click the "CREATE NEW ORGANISM" link at the top of the view, shown in Figure 42. 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 in Figure 43.

	•		PREFERENCES	ENGLISH	•	Close
Name*		Description				
				Canc	el	Save

Figure 43: Create new organism Step 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 44:

WIPO Sequence PR0JECTS	PERSONS & ORGANIZATIONS	ORGANISMS	HELP	-	PREFI	ERENCES ENGLISH 🔻	
ORGANISMS				EXPORT CUSTOM ORGANISMS	IMPORT CUSTOM ORGANISMS	CREATE NEW ORGANISM	1
					Search organism by name	Q	
Name :			De	escription			
test organism			De	escription example			
<u>Demo Organism</u>			De	escription Demo example			
B			bb	b			

Figure 44: Export custom organisms Step 1

Next, a dialog box will open allowing the user to choose the name of the file and the desired file location, see Figure 45.

WIPO Sequence	PROJECTS PERSONS & ORGANIZATIONS ORGANISMS HELP	-	PREFERENCES ENGLISH
ORGANIS	SMS	EXPORT CUSTOM ORGANISMS	IMPORT CUSTOM ORGANISMS CREATE NEW ORGANI
	Save As	×	Search organism by name
Name :	← → ↑ → This PC → Documents → wipo docs ∨ Ŏ	Ø Search wipo docs IIII ▼ ⑦	
<u>test organism</u>	Organise Name Date modified	∏≣≣ ~ (2) Type Size	
<u>Demo Organism</u> <u>B</u>			
Δ	🧼 N-		
	File name: Custom, Organisms, 1625665412516.txt		
	Save as type: txt (*.txt)		
	∧ Hide Folders	Save Cancel	

Figure 45: Export custom organisms, dialog

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 46).

page 40

Custom_Organisms_1625665412516.txt - Notepad File Edit Format View Help A: aaa B: bbb Demo Organism: Description Demo example test organism: Description example Figure 46: Example custom organisms text file

Import Custom Organisms

WIPO ST26 Tool PROJECTS PERSONS	GORGANIZATIONS ORGANISMS	PREFERENCES ENGLISH 🔫
	EXPORT CUSTOM ORGANISMS	IMPORT CUSTOM ORGANISMS CREATE NEW ORGANISM
ORGANISMS		
		Search organism by name Q
Name o		

Figure 47: Import custom organisms Step 1

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 47. This will open an Overlay below the Organisms Table, as shown in Figure 43.

WIPO ST26 Tool PROJECTS PERSON	# Open						×
	🗧 🚽 👻 🛧 📙 > This PC	→ Local Disk (C:) → temp → st2i	i⇒ inbox ·	Search inbox			Q
	Organize 👻 New folder						0
	Outlook	^	Name	Date modified		Туре	
Name :	Scan Shared Temp Trabajo This PC 3D Objects Desktop Bocuments Documents		Custom_Organisms_155557	od 06/03/19 23:19			
Prototype C12	Music						
Demo Organism	Videos		c		-		,
	and the second	Custom_Organisms_155551250078	4.bet	> .bxt (*.bxt) Open		Cancel	~
organisms2019.txt				U	pload	file (.txt	1 (
			Cancel	Import Cu	stom (Organis	ms

Figure 48: Import custom organisms, dialog

- 1) Click on the "Upload file [.txt]" button (1) shown in Figure 48
- 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.

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).

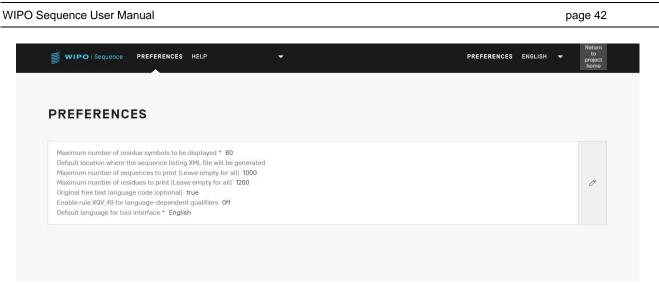


Figure 49: Summary of system preferences

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 49:

WIPO Sequence PREFERENCES HELP	•	PREFERENCES EN	IGLISH v to to proje hore
Maximum number of residue symbols to be displayed * 60			
Default location where the sequence listing XML file will be genera	ated		Select folder
Maximum number of sequences to print [Leave empty for all]			
WARNING: If the project to print contains a large number of sequen	nces, the PDF file may not be generated as t	he resulting file would be too large to view.	
Maximum number of residues to print [Leave empty for all] 2500			
WARNING: If the project to print contains a large number of residue	es within a sequence, the PDF file may not b	e generated as the resulting file would be too larg	je to view.
Original free text language code (optional)			
Enable rule XQV_49 for language-dependent qualifiers			
Default language for tool interface * English			v

Figure 50: 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 60residues.
- 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 whenWIPO Sequence is launched. By default, this is English.

Note:

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

PROJECT DETAILS

This section details the functionalities provided in the Project Details View.

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 51.



Figure 51: Print project Step 1

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 52).

WIPO Sequence DEMO VERIFICATION LANGUAGE REPORT QUALIFIERS	IMPORT DISPLAY THE HELP
GENERAL INFORMATION SEQUENCES	
DEMO	Print Export Import Another Project Validate Generate Sequence Listing
 Print General Information Part Print Sequences 	
Sequence lds 1-12	
Use ',' to separate your ID selection: to select a range use '-' between the empty to select 'ALL'.	the ids or leave
	Cancel Print

Figure 52: Print project Step 2

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 Figure 9.

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 following the steps described in section following in Figure 53.

WIPO Sequence	PROJECT Demo	VERIFICATION REPORT	LANGUAGE DEPENDENT QUALIFIERS	IMPORT REPORT	DISPLAY THE SEQUENCE LISTING	HELP 👻	PREFERENC	ES EI ▼	Return to project home
GENERAL INFORMATION	SEQUENC	ES							
PROJECT	г	Print	1 xport In	nport Anoth	er Project	Validate	Generate	Sequence Lis	ting
			~ -	→ ~ ↑ 🖡	« Docu » se	quences JDD	ڻ ~	,∽ Search sequ	iences JDD
Project Name project der	mo		Orga	nise 🔹 Nev	r folder				III • 🕜
Status invalid Description Original free text languag Automatically add a trans created Off		fier when a CDS fea	ature is	Saved Pictures Screenpresso WIP VTT DATA EME/ File name: Save as type:	viproject demo_16	D-sequence-valid c_163480987199 41987266595	94	10/6/	modified 2021 9:28 AM /2021 11:58 AM > ~ ~ ~
GENERAL INF	FORM	ATION	∧ Hide	e Folders			[Save	Cancel

Figure 53: Export project, dialog

- 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 54.

WIPO Sequence	PROJECT Demo	VERIFICATION REPORT	LANGUAGE DEPENDENT QUALIFIERS	IMPORT REPORT	DISPLAY THE SEQUENCE LISTING	HELP 🔻	PREFERENCES	EIT	Return to project home
SUCCESS: The projec demo_1641987266595		exported succ	essfully at 'C:	\Users\		Documen	ts\sequences J	DD\project	×
GENERAL INFORMATION	SEQUENC	ES							
PROJEC	т	Print	Export	port Anoth	er Project	Validate	Generate Se	quence Listing	3
Project Name project d e	emo		Cr	reation dat	e 2022-01-12				
			Figure 54: Exp	port proje	ct, success				

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 55.

WIPO Sequence	PROJECT	VERIFICATION REPORT	LANGUAGE DEPENDENT QUALIFIERS	IMPORT REPORT	DISPLAY The Sequence Listing	HELP 🔻	PREFERENCES	EI▼	Return to project home
GENERAL INFORMATION	SEQUENC	ES							
PROJEC	г	Print	Export	nport Anoth	er Project	Validate	Generate Sec	uence Listii	ng

Figure 55: Import another project Step 1

The tool will open an Overlay, as shown in Figure 56. The user must first select the project from which they wish to import information.

BASIC	Print Export Import Another Project Validate Generate Sequence Listing
Project*	*
Select the general information contents to be imported	
Select the sequences to be imported. If the box is left unchecked then no sequences will be imported.	
	Cancel Import Project

Figure 56: Import another project Step 2

Next, as shown in Figure 57, 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.

	IMPORT DISPLAY THE HELP
PROJECT DEMO	Print Export Import Another Project Validate Generate Sequence Listing
Project* 1941-AMBIGUOUS	v
9 Select the general information contents to be imported	
2 𝔗 Select the sequences to be imported. If the box is left unchecked then no sequences	s will be imported.
Select Range of Sequence IDs. 1-6	Total Sequences 6
Use commas to separate individual sequences [for example: "2,5,8"]; use a dash to indica	ate a range of sequences (for example: "2-8").
	Cancel Import Project

Figure 57: Import another project Step 3

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 secondary project's General Information, as shown in Figure 58.

Selec	t the general information contents to be imported		
	Element	Origin Element Value	Target Element Value
•	Application Identification	IP Office = FR Application number = 123123123 Filing date = 2022-04-18	IP Office = ES Application number = 1 Filing date = 2021-03-01
1	Applicant File Reference	Applicant file reference = 123124	Applicant file reference = ABC1234
•	Earliest Priority Application Identification	IP Office = FR Application number = 123123 Filing date = 2022-04-18	IP Office = ES Application number = 1 Filing date = 2021-03-01
•	Applicant Name	Name = AP-HM Language code = fr Name Latin = AP-HM	Name = Steven Language code = es
•	Inventor Name		Name = Steven Language code = es
1	Applicant List	Applicant name = AP-HM	Applicant name = Steven
1	Invention Title Bag	Invention title = TEST, Language code = fr	Invention title = AMBIGUOUS, Language code = es

Figure 58: Import Another project Step 4

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 58.

SUCCESS: The project has been imported successfully.

Figure 59: Import another project, success

As shown in Figure 59, a blue banner appears if the elementshave been imported correctly.

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 60.

Sequence User Manu	al								page	e 49
		VERIFICATION REPORT	LANGUAGE DEPENDENT QUALIFIERS	IMPORT REPORT	DISPLAY THE SEQUENCE LISTING	HELP	•	PREFERENCES	EI▼	Return to projec home
GENERAL INFORMATION	SEQUENCES	5								
PROJECT	·	Print	Export In	nport Anoth	er Project	Valida	ate	Generate Sec	quence Lis	ting
							_			
Project Name project dem	0				∋ 2022-01-12					
Project Name project dem Status new	0			Creation dat File Name	∋ 2022-01-12					
A DESCRIPTION OF A DESC	0		F							ß
Status new			F	File Name Sequences 2		uage cod	le			Ø

Figure 60: Project validation Step 1

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 61 shows the screen that will be displayed in the case of successful validation.

w	IPO Sequence	TESTSTRAIN	VERIFICATION REPORT	LANGUAGE DEPENDENT QUALIFIERS	IMPORT Report	DISPLAY THE SEQUENCE LISTING	HELP▼	PREFERENCES	[▼	Return to project home
SUCCE	SS: The project	has been sud	ccessfully veri	fied						×
Rej	port successfully ge	enerated on2022	-01-12							
Thi	is project does not c	contain any error	s or warnings.							

Figure 61: Successful project validation

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 62. Each row identifies whether this is an error, which must beaddressed, or a warning, which can be ignored by the user.

Seque	ence User Man	ual								pag	ge 50
w	IPO Sequence	PROJECT Demo	VERIFICATION REPORT	LANGUAGE DEPENDENT QUALIFIERS	IMPORT Report	DISPLAY THE SEQUENCE LISTING	HELP	•	PREFERENCES	EI▼	Return to project home
RROR:	: After project v	/erificatior	n, some errors	or warnings	have beer	n detected.					×
Repo	ort genera	nted on	2022-01	-12			Delete	select	ed sequences	Print	Report
?epc	ort genera	ited on	0 2022-01	-12			Delete	select	ed sequences	Print	Report
Repo	D rt genera	ited on Data Ele		-12 Message Te	ext	Detecte			ted sequences		
•	-		ement o			Detecte					

Figure 62: Project validation, error/warnings

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).

GENERAL INFORMATION SEQUENCES PROJECT Print Export Import Another Project Validate Generate Sequence Listing Project Name project demo Creation date 2022-01-12 Status invalid File Name Sequences 2 Non English free text language code Validate Import Another Project Import Another Project	WIPO Sequence	PROJECT Demo	VERIFICATION REPORT	LANGUAGE DEPENDENT QUALIFIERS	IMPORT REPORT	DISPLAY THE SEQUENCE LISTING	HELP	•	PREFERENCES	EI T	Return to project home
Project Name project demo Creation date 2022-01-12 Status invalid File Name Description Sequences 2	GENERAL INFORMATION	SEQUENCI	ES								
Status invalid File Name Description Sequences 2	PROJECT		Print	Export	nport Anoth	er Project	Valid	ate	Generate Sec	uence Li	sting
Automatically add a translation qualifier when a CDS feature is	Status i nvalid Description Original free text language	e code	ijer when a CDS fe	F	ile Name equences :	2	lage co	de			Ø

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.

DEMO	Print Export Import Another Project Validate Generate Sequence Listing
	Save As
Project Name Demo Status valid	← → ^ ▲ Image: Search DEMO Organise ▼ New folder Image: Search DEMO
Description Original free text language code de Automatically add a translation qualifier when a Cl	This PC Name Status Date modified Desktop Documents Documents Downloads Movice
GENERAL INFORMATION	I make
SEQUENCES	File name: Demoxml V Save as type: XML (*xml) V A Hide Folders Save Cancel

Figure 64: Generate sequence listing, dialog

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.

E	RROR:	The ST.26 sequ	uence listing file cannot	be generated because	the pro	ject contains er	rors.	×
R	Repo	ort genera	ted on 2022-01	-21		Delete selected	sequences Print Repo	ort
		Severity o	Data Element o	Message Text	Detec	ted Value	Detected Sequence	
		ERROR	<u>Qualifier Molecule</u> <u>Type</u>	The value of the qualifier 'mol_type' is not one of the permitted values for a DNA sequence or the value is missing.	rRNA		Sequence 1	
		ERROR	Feature Location	The feature location includes a residue number greater than the length of the sequence, which is invalid.	150		Sequence 1	

Figure 65: Generate sequence listing, errors/warnings

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.

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.

PLICATION IDENTIFICATION		
		ſ
Application Identified Before the assignment of the application number	IP Office AU - Australia	

Figure 66: Edit application identification

To edit information within the Application Identification subsection, click on the pencil icon highlighted above, 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.
- 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.

BD D		Ψ.
Application Identification*	Applicant file reference* 123456 3 Application File Reference is a mandatory field when 'Before the assi number' is provided.	gnment of the application
Application number 12345	Filing date 2019-02-20 5	Ĩ

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."

Priority Identification

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

page 53

pplication Identified Before the assignment of the application number	IP Office AU - Australia	
pplicant file reference 123	Filing date 2022-01-07	C

Figure 68: Edit/add priority application Step 1

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 Priority Application" button in the Overlay, shown in Figure 69.

PRIORITY IDENTIFICATION Add Priority Application			
IP Office*	•	Filing date Date	
Application number*		Selected Earliest Priority Application*	Ŧ
		Cancel Add Priority Appli	

Figure 69: Priority application edit panel

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 <u>also</u> 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".

Add Applicant	Add Inventor		
Select the option* • Existing applicant • New applicant			
Select applicant*			•
הילה בן אברהם Juan Rodriguez Garc John Smith	ia		

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 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.

Select the option* Existing applicant New applicant 		
Name*		Residence Address
Language*	Ŧ	Correspondence Address
Name Latin		Mark as primary
		Cancel Add Applicant

Figure 71: Add new applicant/inventor

Invention Title

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

 To add a new invention title, click on the "Add Invention title" button, as shown in Figure 72.

page 55

page	56

INVENTION TITLE		
Add Invention title		
Invention title*	Language*	Ŧ
ß	Cancel Add Invention	title



- 8) In this Overlay, the user must enter the title of the invention and also indicate what language the title is provided in.
- 9) Click the blue "Add Invention title" button.

Note: 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.

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.

Create Sequence

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

WIPO S	equence User Manual					page 57
GEI	NERAL INFORMATION SE	QUENCES				
- S	EQUENCES					
	Create new sequence	Import sequence	Insert Sequence	e Reorder Sequ	ence Bulk Edit	
	Sequence ID Number	Sequence Name	Length	Molecule Type	Organism	Skipped Sequence
	1	DNA_RNA_withut_ ST25_seq_1	10	DNA	Homo sapiens	No

Figure 73: Create new sequence Step 1

The remainder 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"⁵ 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. The user simply starts typing the name of the organism and a dropdown list will appear with options from both of these sources (4).

The user can also classify the organism by selecting a Qualifier Molecule Type⁶ 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

⁵ See WIPO Standard ST.26 – Annex I, Sections 1 & 3 for the tables of valid Nucleotide symbols & Amino acid symbols

⁶ 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

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.

Sequence Name Name/Description for sequence	1		Molecule Type* DNA	2		,
tesidues" cractularty adutactularty adutactulart to attgaat ctactto attgaat ctactto attgaat gaat ctactto attgaat ctactto attgaat ctactto tractto attgaat ctactto attgaat ctactto attgaat ctactto attgaat ctactto attgaat pact cactto attgaat ctactto attgaat ctattgaat ctactto attgaat ctactto to attgaat ctactto attgaat ctactto attgaat ctactto attgaat ctactto attgaat ctactto attgaat ctactto attgaat pact actto attgaat ctactto attgaat ctactto attgaat ctactto attgaat ctactto attgaat ctactto	ctacttcattgaatc tcattgaatctacttu gaatctacttcattga tacttcattgaatcta cattgaatctacttc gaatctacttcattga tacttcattgaatct	tacttcattgaatctactto cattgaatctacttcattg aatctacttcattgaatcta acttcattgaatctactto attgaatctacttcattga aatctacttcattgaatct	cattgaatctacttcattgaatctac aatctacttcattgaatctacttcat acttcattgaatctacttcattgaat attgaatctacttcattgaatctact alstcacttcattgaatctacttcatt acttcattgaatctacttcattgaat	ttcattgaatctacttcattgaat tgaatctacttcattgaatctac ctacttcattgaatctacttcatt tcattgaatctacttcattgaatc gaatctacttcattgaatctactt ctacttcattgaatctacttcattg	ctacttcattgaatctacttcattgaat tcattgaatctacttcattgaatctact gaatctacttcattgaatctacttcatt tacttcattgaatctacttcattgaat cattgaatctacttcattgaatctactt gaatctacttcattgaatctactt	ctacttcat ttcattgaat gaatctact ctacttcatt tcattgaat gaatctact
Organism name *			Qualifier Molecule	Туре		
	-					
Saaristoa firma	4		genomic DNA		5	
Mark as an intentionally skipped sequenc The sequence contains both DNA & RNA fr	ragments	B of the DNA and RNA fra			5	
Saaristoa firma Mark as an intentionally skipped sequence The sequence contains both DNA & RNA fr feature with the key 'misc_feature' is recom folecule Type	ragments	-		Further Text	5	

Figure 74: Create new sequence edit panel

 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. An example of a newly created sequence is shown in Figure 75.

For details on how to reorder the Sequence list, see Figure 85 and Figure 86.

		1 0
SEQUENCE 2		
Sequence Number (ID) 2 Sequence Name Name/Description for sequence Length 1792	Molecule Type DNA Organism Saaristoa fi Contains DNA and RNA	
FEATURES		
Feature Key	Location	Qualifiers
misc_feature	713	note = RNA 🖉
source	11792	moLtype = genomic DNA organism = Saaristoa firma
	₩ 4 1 ≫ ₩	
SEQUENCE		
atctactica tigaatctac ticatigaat ctactica	t gaatctactt cattgaatct acttcattga atctac	tca 80 ^
ttgaatctac ttcattgaat ctacttcatt gaatctac	t cattgaatct acttcattga atctacttca ttgaat	
ttcattgaat ctacttcatt gaatctactt cattgaat		
ctacttcatt gaatctactt cattgaatct acttcatt		
gaatctactt cattgaatct acttcattga atctactt	a ttgaatctac ttcattgaat ctacttcatt gaatcta	actt 400



Import Sequence

Sequences can also be imported directly from files into a project. The accepted file formats⁷ are **raw, multi-sequence, FASTA, ST.26 and ST.25.** When selected, WIPO Sequence will automatically detect the format used in the file.

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

EQUENCES				
Import sequence	Insert Sequence	Reorder Sequ	ence Bulk Edit	
Sequence Name	Length	Molecule Type	Organism	Skipped Sequence
DNA_RNA_withut_ ST25_seq_1	10	DNA	Homo sapiens	No
DNA_RNA_withut_ ST25_seq_2	30	RNA	Abrophyllum ornans	No
	Import sequence Sequence Name DNA_RNA_withut_ ST25_seq_1 DNA_RNA_withut_	Import sequence Insert Sequence Sequence Name Length DNA_RNA_withut_ 10 DNA_RNA_withut_ 30	Import sequence Insert Sequence Reorder Sequence Sequence Name Length Molecule Type DNA_RNA_withut_ 10 DNA DNA_RNA_withut_ 30 PNA	Import sequence Insert Sequence Reorder Sequence Bulk Edit Sequence Name Length Molecule Type Organism DNA_RNA_withut_ 10 DNA Homo sapiens DNA_RNA_withut_ 30 PNA Abrophyllum

Figure 76: Import sequence Step 1

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.

page 59

⁷ See File Formats section of this document

					Upload file [.txt, .xml]
	Mabrir			×	
	-				Cancel Import sequence
		cumentos > Import Seq Listing	v ひ Buscar en Impor	rt Seq Listing 🔎	
	Organizar 👻 Nueva ca	arpeta	l	88 - 🔟 🕜	
	everis ^	Nombre	Fecha de modifica	Tipo ^	
	ST26 authoring t	ST26T-1087	02/07/2019 13:54	Carpeta de archiv	
		9_ST25_9238084_DTD_Based.xml	17/08/2017 18:33	Documento XML	
EQUENCE 1	Este equipo	🐑 changedata.xml	04/07/2019 10:17	Documento XML	
	👆 Descargas	Insecticidal Proteins (1).xml	05/07/2019 12:52	Documento XML	
	Documentos	Insecticidal Proteins.xml	25/06/2019 18:50	Documento XML	
Sequence Number (ID) 1	Escritorio	Invalid_project.xml	17/07/2019 12:07	Documento XML	
Sequence Name US20190136258A1-2019050	📰 Imágenes	🔮 Noseqs0.xml	08/07/2019 14:44	Documento XML	O
Length 245	👌 Música	Opt_Soy_Loci.xml	11/03/2019 15:45	Documento XML	
	Objetos 3D	QN_1.xml Results QV 29.xml	02/07/2019 14:37 11/07/2019 15:00	Documento XML Documento XML	
	Vídeos	ST_25_import_complex.txt	05/07/2019 13:42	Documento ANL Documento de te	
	OSDisk (C:)	ST_25_long.txt	04/07/2019 14:24	Documento de te	
EATURES			0401720101424	>	
		bre: ST_25_import_complex.txt	 .txt .xml (*.txt;*.; 	vm) ×	
	NOIL	are areas import complexity	· · · · · · · · · · · · · · · · · · ·		

Figure 77: Import sequence, dialog

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 last section ofthis 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 of Sequence IDs" 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 sequences 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".

Sequence Number [ID]		Position	
[1]		1	
[2]		2	
[3]		3	
	₩ 4 1	Ж	
Select Range of Sequence IDs. 1-3			

Figure 78: Import sequence, ST.25/ST.26 format

In the case of importing a multi-sequence format file (see Figure 79), the user will see a "Select Range of Sequence IDs" 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.

O Sequence User Manual	page 62
C:\Users\ Downloads\multisequence_u_t.txt	Upload file [.txt, .xml, .FASTA]
Enter the sequences to be imported	
Sequence Number [ID]	Position
[1]	< First Sequence;DNA;Abies alba>
[2]	< Second Sequence;RNA; >
[3]	< Second Sequence;AA; >
[4]	< Third Sequence; DNA; Abies alba>
ka a 1	
Select Range of Sequence IDs.	
Use commas to separate individual sequences (for example: "2,5,8"); use a dash to indicate a range of sequences (for example: "2-8").	

Figure 79: Import sequence, multi-sequence format

The last two formats that are accepted by the import sequence process are the raw and FASTA file formats (see the section on file formats). These formats 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 andwhen a FASTA file is imported Edit panel shown in Figure 81 is displayed. The user should proceed by providing the mandatory fields.

:\Users\	\3AAraw.txt	Upload file [.txt, .xml, .FASTA]
he file selected is in raw format. Please select th	e sequence molecule type in order to import the sequence.	
Sequence Name	Molecule Type*	
Organism name *	Qualifier Molecule Type	

Figure 80: Import sequence, raw format

page 63

C:\Users\#	.FASTA\IDNAsequence.fasta Upload file [.bxt, .xml, .FASTA				
Sequence Number [ID]	Description				
1	>HM118516.1 Uncultured bacterium hypothetical protein gene, partial cds; hypothetical protein and Est1 [est1] genes, complete cds; and putative DNA polymerase I [polA] gene, partial cds				
14 4 1	Э Э				
Sequence Number (ID) 1	Sequence Name				
Molecule Type*	v				
lease select the sequence molecule type for the sequence that you would like to import.					
Organism name*	Qualifier Molecule Type*				
Check to save description as a note.					
	Cancel Import sequence				

Figure 81: Import sequence, FASTA 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, shown in Figure 127.

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).

Create new sequence	nport sequence Insert Sequen	ce Reorder S	equence Bulk Edit		
Sequence ID Number	Sequence Name	Length	Molecule Type	Organism	Skipped Sequence
l	QV_31K_Test_File_seq_2	20	AA	Saccharomyces cerevisiae	Yes
2	QV_31K_Test_File_seq_3	90	DNA	Saccharomyces cerevisiae	Yes
3	QV_31K_Test_File_seq_4	20	AA	Saccharomyces cerevisiae	Yes
<u>1</u>	QV_31K_Test_File_seq_5	90	DNA	Saccharomyces cerevisiae	Yes
2	QV_31K_Test_File_seq_6	20	AA	Saccharomyces cerevisiae	Yes

page 64

An Overlay with a panel will appear (see Figure 83). The user must fill out all the information required for creating a sequence (see Figure 73), 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).

Sequence Number (ID)* 3		
Sequence Name	Molecule Type*	
inserted sequence	DNA	
ytaggaatgtaggaatgtaggaatctacatgtaggaatgtaggaatgtaggaatctacatgtaggaatgtaggaa gtaggaatgtaggaatgtaggaatctacatgtaggaatgtaggaatgtaggaatctacat	tgtaggaatctacatgtaggaatgtaggaatgtaggaatctacatgtaggaatgtaggaatgtaggaa	tctaca
Organism name *	Qualifier Molecule Type	
	Qualifier Molecule Type genomic DNA	,
Organism name * Wohlfahrtiopsis bishoppi Mark as an intentionally skipped sequence The sequence contains both DNA & RNA fragments		

Figure 83: Insert sequence edit panel

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

Create new sequence Insert Sequence Reorder Sequence Bulk Edit							
Sequence Name	Length	Molecule Type	Organism	Skipped Sequence			
QV_31K_Test_File_seq_2	20	AA	Saccharomyces cerevisiae	Yes			
QV_31K_Test_File_seq_3	90	DNA	Saccharomyces cerevisiae	Yes			
inserted sequence	210	DNA	Wohlfahrtiopsis bishoppi	No			
QV_31K_Test_File_seq_4	20	AA	Saccharomyces cerevisiae	Yes			
QV_31K_Test_File_seq_5	90	DNA	Saccharomyces cerevisiae	Yes			
	Sequence Name QV_31K_Test_File_seq_2 QV_31K_Test_File_seq_3 inserted sequence QV_31K_Test_File_seq_4	Sequence Name Length QV_31K_Test_File_seq_2 20 QV_31K_Test_File_seq_3 90 inserted sequence 210 QV_31K_Test_File_seq_4 20	Sequence Name Length Molecule Type QV_31K_Test_File_seq_2 20 AA QV_31K_Test_File_seq_3 90 DNA inserted sequence 210 DNA QV_31K_Test_File_seq_4 20 AA	Sequence Name Length Molecule Type Organism QV_31K_Test_File_seq_2 20 AA Saccharomyces cerevisiae QV_31K_Test_File_seq_3 90 DNA Saccharomyces cerevisiae inserted sequence 210 DNA Wohlfahrtiopsis bishoppi QV_31K_Test_File_seq_4 20 AA Saccharomyces cerevisiae QV_31K_Test_File_seq_5 90 DNA Saccharomyces cerevisiae			

Figure 84: Sequence inserted in specified position

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.

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.

Create new sequence	nport sequence Insert Sequen	ce Reorder S	Bulk Edit		
Sequence ID Number	Sequence Name	Length	Molecule Type	Organism	Skipped Sequence
1	QV_31K_Test_File_seq_2	20	AA	Saccharomyces cerevisiae	Yes
2	QV_31K_Test_File_seq_3	90	DNA	Saccharomyces cerevisiae	Yes
3	inserted sequence	256	DNA	Wohlfahrtiopsis bishoppi	No
4	QV_31K_Test_File_seq_4	20	AA	Saccharomyces cerevisiae	Yes
5	QV_31K_Test_File_seq_5	90	DNA	Saccharomyces cerevisiae	Yes
<u>6</u>	QV_31K_Test_File_seq_6	20	AA	Saccharomyces cerevisiae	Yes

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.

<u>10</u> 2019050 S00001_)9- 1458 seq_SEQ ID NO 15	8	Artificial sequence
	M 1 2 3	3 4 5 ≫ ₩	
Sequence Number (ID)* 3 To select a range use "-" between the ids		New Sequence Position 2	
			Cancel Reorder Sequence

Figure 86: Reoder sequence edit panel

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

•	SEQUENCES Create new sequence	sequence	e Reorder Seque	nce Bulk Edit		
	Sequence ID Number	Sequence Name	Length	Molecule Type	Organism	Skipped Sequence
	1	QV_31K_Test_File_seq_2	20	AA	Saccharomyces cerevisiae	Yes
	2	inserted sequence	256	DNA	Wohlfahrtiopsis bishoppi	No
	3	QV_31K_Test_File_seq_3	90	DNA	Saccharomyces cerevisiae	Yes
	4	QV_31K_Test_File_seq_4	20	AA	Saccharomyces cerevisiae	Yes
	5	QV_31K_Test_File_seq_5	90	DNA	Saccharomyces cerevisiae	Yes

Figure 87: Sequence/s shown as reordered

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 88:

I	GENERAL INFORMATION S	EQUENCES				
	SEQUENCES					
	Create new sequence	Import sequence	Insert Sequence	e Reorder Sequ	ence Bulk Edit	
	Sequence ID Number	Sequence Name	Length	Molecule Type	Organism	Skipped Sequence
	1	DNA_RNA_withut_ ST25_seq_2	30	RNA	Abrophyllum ornans	No
	2	DNA_RNA_withut_ ST25_seq_1	10	DNA	Homo sapiens	No

Figure 88: Bulk edit Step 1

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

Type of bulk edit* Qualifier molecule type	Ŧ	Select Range of Sequence IDs*		
Select the type of bulk edit to continue		Use commas to separate individual sequences [for example: "2,5,8"]; use a dash to indicate a range of sequences [for example "2-8"].		
Molecule Type*	Ŧ	Qualifier Molecule Type*		
Only DNA/RNA sequences allow editing of the qualifier 'mol_typ because for amino acid sequences this qualifier value is automatically set to 'protein'.)e'			
The qualifier 'mol_type' of hybrid DNA/RNA sequences, where the 'comb	pinedIndic	- :ator' = 'Yes', cannot be bulk edited.		
Enter one or more of the following DNA sequences to edit the qualifier 'r	mol_type'.			

Figure 89: Bulk edit Step 2

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

Type of bulk edit* Qualifier molecule type		Ŧ	Select Range of Sequence IDs* 3,7					
Select the type of bulk edit to continue					separate individual s ash to indicate a rang		ences (for example: sequences (for exam	ple:
Molecule Type*				Qualifier Molecu	le Type*			
DNA		*	unassigned DI	NA			Ŧ	
Only DNA/RNA sequences allow editing of the qualifier 'mol_type' because for amino acid sequences this qualifier value is automatically set to 'protein'.								
The qualifier 'mol_type' of hy	ybrid DNA/RNA sequences, whe	re the 'combir	nedIndic	ator' = 'Yes', cannot b	e bulk edited.			
Enter one or more of the foll	owing DNA sequences to edit th	ne qualifier 'm	ol_type'.					
Sequence ID Sequence Name Length				Molecule Type	Organism		Qualifier molecul type	e
3	500 sequences_seq_3	15		DNA	Vaccaria hispanica		genomic DNA	

Figure 90: Bulk edit, qualifier mol_type identified

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: Bulk edit, updated qualifier mol_type

When the type of bulk edit is Organism:

 Choose "Type of bulk edit" as 'Organism' (1) as shown in Figure 92. The user must enter the range of SEQ IDs to be edited (2). Then if for example the user has chosen to modify the value of organism to "synthetic construct", the system will notify him that the qualifier 'mol_type' will be automatically changed to "other DNA" or "other RNA" according to molecule type(3).

Type of bulk edit* Organism			Ŧ	Select Range of Se	quence IDs*		
Select the type of bulk edit to c	ontinue		Use commas to separate individual sequences (for example: "2,5,8"); use a dash to indicate a range of sequences (for example: "2-8").				
Organism name* synthetic construct The qualifier 'mol_type' value will be automatically changed to 'other DNA' or 'other RNA'.							
Sequence ID Number	Sequence Name	Length	M	olecule Type	Organism	Qualifier molecule type	
1	Seq_1	192	DI	AA	Mus musculus	genomic DNA	
2	Seq_2	20	DI	A	Homo sapiens	other DNA	
3	Seq_3	20	DM	A	Gabaza connectens	other DNA	
144 × 1							

Figure 92: Bulk edit, organism Step 1

When the type of Bulk edit is Features:

1) Choose "Type of bulk edit" as 'Features' (see Figure 93) (1). Then the user need to specify if they want to edit existing features or add new ones (2). The user must enter the 'Molecule Type' (3) and the range of sequence IDS to be edited (4). In the case where the user selects to "Edit feature' (2), then if the user has chosen to modify the value of the feature location (6) to "complement(join(1..30,61..90))", the tool will update all existing CDS features (5) for SEQ ID 5 and 8 to the provided value.

Sequence User Man	ual					page 69	
		K 	2	3 4 5 🏓	₩		
Type of bulk edit* Features			Ŧ	Select Range of Sa 5,8	tence IDs*		
Select the type of bulk edit to continue				Use commas to sep indicate a range of s	arate individual sequences (fo sequences (for example: "2-8"	r example: "2,5,8"]; use a dash t].	0
Type of bulk feature edit* Edit feature	2			Molecule Type* RNA	3		Ŧ
Feature Key* CDS	6		~	Feature Location* 112	6		
For a detailed explanation of the f	ormat of the feature location please	click here					
Sequence ID Number	Sequence Name	Length		Molecule Type	Organism	Qualifier molecule type	
5	40000 sequences_seq_5	16		RNA	Vaccaria hispanica	genomic RNA	
8	40000 sequences_seq_8	16	I	RNA	Gabara gigantea	genomic RNA	
10	40000 sequences seq 10	15	1	RNA	Wabasso hilairoides	genomic RNA	

Figure 93: Bulk edit, feature step 1

2) In the case where the user selects the type of bulk feature edit as "Add feature" (see Figure 94): a new feature "CHAIN" (4). with feature location "1..13" (5) will be added to each of the selected SEQ IDs 1,2,7 (3).

Type of bulk edit* Features			Select Range of Sequence IDs* 1.2.7 3			
ntinue			Use commas to sep indicate a range of	arate individual sequences (fo sequences (for example: "2-8'	or example: "2,5,8"]; use a dash t '].	0
Type of bulk feature edit* Add feature				2		Ŧ
4		Ŧ	Feature Location* 113	5		
at of the feature location please <u>cl</u>	ick here					
Sequence Name	Length	1	Molecule Type	Organism	Qualifier molecule type	
40000 sequences_seq_1	15	ļ	AA	Gabara gigantea	protein	
40000 sequences_seq_2	16	ļ	AA	Daboia russelii	protein	
40000 sequences_seq_7	15	A	AA	Wabasso hilairoides	protein	
	Sequence Name 40000 sequences_seq_1 40000 sequences_seq_2 40000	Image: I	Image: sequences_seq_2 16	1.27 3 ntinue Use commas to sep indicate a range of section and the feature location please click here 4 Feature Location* 1.13 Feature Location* 1.13 AA 40000 15 40000 15	12.7 3 ntinue Use commas to separate individual sequences [frindicate a range of sequences [for example: "2-8" indicate a range of sequences [for example: "2-8" AA 1 Y Ad 2 4 Y Feature Location* 113 3 Feature Location* 113 Sequence Name Length Molecule Type Organism 40000 sequences_seq_1 15 AA Gabara gigantea 40000 sequences_seq_2 16 AA Daboia russelii	12.7 3 ntinue Use commas to separate individual sequences (for example: "2.5.8"); use a dash tindicate a range of sequences (for example: "2.6.8"); use a dash tindicate a range of sequences (for example: "2.6.8"); use a dash tindicate a range of sequences (for example: "2.6.8"); use a dash tindicate a range of sequences (for example: "2.6.8"); use a dash tindicate a range of sequences (for example: "2.6.8"); use a dash tindicate a range of sequences (for example: "2.6.8"); use a dash tindicate a range of sequences (for example: "2.6.8"); use a dash tindicate a range of sequences (for example: "2.6.8"); use a dash tindicate a range of sequences (for example: "2.6.8"); use a dash tindicate a range of sequences (for example: "2.6.8"); use a dash tindicate a range of sequences (for example: "2.6.8"); use a dash tindicate a range of sequences (for example: "2.6.8"); use a dash tindicate a range of sequences (for example: "2.6.8"); use a dash tindicate a range of sequences (for example: "2.6.8"); use a dash tindicate a range of sequences (for example: "2.6.8"); use a dash tindicate a range of sequences (for example: "2.6.8"); use a dash tindicate a range of sequences (for example: "2.6.8"); use a dash tindicate a range of sequences (for example: "2.6.8"); use a dash tindicate a range of sequences (for example: "2.6.8"); use a dash tindicate a range of sequences (for example: "2.6.8"); use a dash tindicate a range of sequences (for example: "2.6.8"); use a dash tindicate a range of sequences (for example: "2.6.8"); use a dash tindicate a range of sequences (for example: "2.6.8"); use a dash tindicate a range of sequences (for example: "2.6.8"); use a dash tindicate a range of sequences (for example: "2.6.8"); use a dash tindicate a range of sequences (for example: "2.6.8"); use a dash tindicate a range of sequences (for example: "2.6.8"); use a dash tindicate a range of sequences (for example: "2.6.8"); use a dash tindicate

Figure 94: Bulk edit features: add feature

3) In the case where the user selects type of bulk edit "Bulk Skip" (see Figure 95). The range of Sequences to be skipped is identified as a range of SEQ IDs (1).

	M	1 🔅
Type of bulk edit* Bulk skip	v	Select Range of Sequence IDs* 1
Select the type of bulk edit to continue		Use commas to separate individual sequences (for example: "2,5,8"); use a dash to indicate a range of sequences (for example: "2-8").

Figure 95: Bulk edit features: bulk skip

Edit Sequences

To edit a sequence, the user must click on the Sequence ID Number of the corresponding sequence to be modified, as highlighted in Figure 96(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.

▼	SEQUENCES					
	Create new sequence	Import sequence	Insert Sequence	Reorder Sequ	ence Bulk Edit	
	Sequence ID Number	Sequence Name	Length	Molecule Type	Organism	Skipped Sequence
	1	DNA_RNA_withut_ ST25_seq_2	30	RNA	Abrophyllum ornans	No
	2	DNA_RNA_withut_ ST25_seq_1	10	DNA	Homo sapiens	No

Figure 96: Edit sequence

This will open a new Section in the Project Detail View beneath the Sequences Section, as shown in Figure 97.

This Sequence Section is composed of the "Basic Information" Edit Panel highlighted and shown first, the "Features" listTable shown second and the "Sequence" residues Edit Panel at the end .

0	SENERAL INFORMATION SEQUENCES				
Ŧ	SEQUENCE 1				
	Sequence Number 0] 1 Sequence Name ST25_221_BINDING_edited_seq_1 Length 10	Molecule Type AA Organism synthetic co	nstruct		0
Ŧ	FEATURES				
	Add feeture				
	Feature Key	Location	Qualifiers		
	source	110	mol_type = protein organism = synthetic construct		0
		166 - 66 1 De Del			
Ŧ	SEQUENCE				
	ACXEWWXXXC			10	
					0
			<	1/1 >	

Figure 97: Edit sequence, edit panel

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: <u>x</u>
- Residue numbers delimiting a sequence span: <u>x..y</u>
- Residues before the first or beyond the last specified residue number: <a>x, >x, <x..y, <a>x, >y, <x..>y
- A site between two adjoining nucleotides: <u>x^y</u>
- 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 98.

→ SEQUENCE 2								
	Sequence Number (ID) 2 Sequence Name Inserted sequence Length 41		Molecule Type DNA Organism Wohlfahrtiopsis bishoppi					
•	▼ FEATURES							
	Feature Key	Location	Qualifiers					
	Source	141	mol_type = genomic DNA organism = Wohlfahrtiopsis bishoppi					
Ŧ	* SEQUENCE							
	gatagtatgt atatatagta gtatgatgat gata	itgatga t	41					
		Figure 98: Add feature to a	sequence					

Next, in the Overlay that opens (shown in Figure 99) select an entry from the feature key⁸ dropdown list and specify the feature location within the sequence that the feature applies to.

⁸ See WIPO Standard ST.26 – Annex I, Sections 5 & 7 for tables of feature keys for Nucleotides and Amino acids

Sequence User Mar	nual		page 73
FEATURES			
Add feature			
Feature Key		Qualifiers	
source	123	mol_type = genomic DNA organism = Tinamus osgoodi	÷
		M	
Feature 1			
Feature Key* D_segment		Feature Location* 4.15	
For a detailed explanation of the forma	at of the feature location please click here		

Figure 99: Add feature overlay

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

Feature 1		
Feature Key* CDS	Ŧ	Feature Location* 113
Automatic addition of a translation qualifier is currently on. This means a separate sequence will be generated and a protein_id qualifier created if the translation qualifier value meets the minimum len requirement.		
For a detailed explanation of the format of the feature location please <mark>click here</mark>		

Figure 100: Feature location detailed explanation pop-up link

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 100 to make it easier for the user to identify it.

To finish, click the blue "Create Feature" button, shown in Figure 99.

CDS Features

The CDS Feature type is used to describe the coding sequence for a protein⁹. A CDS feature may optionally include the amino acid translation of the segment of the sequence to which it belongs

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

page 74

and if this satisfied the minimum length requirement will appear as a separate sequencewithin 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.

Note:

From version 2.1.0, the 'Automatically add a translation qualifier...' checkbox is ticked by default

Automatic CDS Feature creation

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

- 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.

QUALIFIERS			
Qualifier Name * translation	Ŧ	Select Genetic Code 1 - Standard Code	Ŧ
Sequence Name		Qualifier Value Automatic Translation is ON. Leave this field blank to automatically generate the transla qualifier value, protein_id qualifier, and associated amino acid sequence.	ation
		Cancel Create Our	alifier

Figure 101: 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 102.

WIP	D Sequence User Manual		page 75	5
•	FEATURES			
	Add feature Feature Key	Location	Qualifiers	
	CDS	113	protein_id = 3 translation = MYIN	Ø
	source	1237	mol_type organism = test bla	Ø
		H (1) H		

Figure 102: New CDS feature in feature table

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 103.

Ŧ	SEQUENCE 3				
	Sequence Number (ID) 3 Sequence Name Seq_3 Length 4		cule Type AA nism Caballerocof	tylo klawei	Ø
v	FEATURES Add feature				
	Feature Key	Location		Qualifiers	
	COILED	14		note = qual val	Ø
	source	14		rno∟type = protein organism = Caballerocotyla klawei	Ø
		100 -00 1	₩ ₩		

Figure 103: CDS feature, generated amino acid 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 3characters).

April 2023

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.

WIPO Sequence PROJECT VERIFICATION LANGU DEPEN REPORT QUAL	NDENT	IMPORT REPORT	DISPLAY THE Sequence Listing	HELP	•	PREFERENCES	EI▼	Return to project home
Project Name* project demo		Creation 2022-01-						
Status modified	v	File Nan	ne					
Description		Sequenc 2	ces					
Original free text language	Ŧ	Non Eng	lish free text la	angu				Ŧ
Automatically add a translation qualifier when a CDS feature created	e is							
Automatically add a translation qualifier, its value, and a protein_id qualif and associated separate amino acid sequence (if appropriate) when a CD feature is added to a nucleotide sequence.								
		-				Can	cel	Save

Figure 104: Automatic generation of translation qualifier checkbox: uncheck to turn-off

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 (if the checkbox identified in Figure 105 is checked) when a pseudo or pseudogene qualifier is added to a CDSfeature, 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 featureand delete the translation and protein_id qualifiers, and then update the feature.

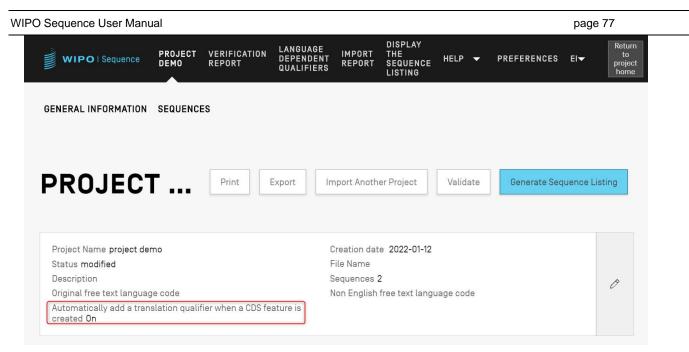


Figure 105: Automatic generation of translation qualifier: ON

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 isno 'transl_except' qualifier indicating that the stop codon is to be translated into a particular aminoacid, as shown in Figure 106 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.

	Jal		page 78
ERROR: No translated sequence w qualifiers 'transl_except' or 'codor	ill be generated as the CDS feature contains an improper in start' are required.	internal stop codon. Please ensure that the correct genetic code table h	has been selected or whether
Add feature			
Feature Key	Location	Qualifiers	
CDS	113	translation	0
source	1237	mol_type organism = test bla	0
	100 40	(1)> M	
SEQUENCE			
SEQUENCE atg <mark>taa</mark> atca acccagtagt ac	stacaatca tgatcaacco agtagtacta caatcatgat c	caacccagta gtactacaat catgatcaac ccagtagtac tacaatcatg	atcaacccag tagt
	tacaatca tgatcaacco agtagtacta caatcatgat (caacccagta gtactacaat catgatcaac ccagtagtac tacaatcatg	atcaacccag tagt
	tacaatca tgatcaacoo agtagtacta caatcatgat (caaccoagta gtactacaat catgatcaac coagtagtac tacaatoatg	atcaacccag tagt

Figure 106: Automatic generation of translation qualifier ON and sequence with 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 107.

w	IPO Sequence	TEST Project	VERIFICATION REPORT	LANGUAGE DEPENDENT QUALIFIERS	IMPORT Report	DISPLAY THE SEQUENCE LISTING	HELP	¥	PREFERENCES	ENGLIS-	Return to project home
ERROR	After project	verification	, some errors c	or warnings have been (detected.						×
Repo	ort genera		2022-02 -	-12 Message	Text	Detec	ted Value		ete selected seque Detected Sequ		nt Report
	ERROR	Ē	Feature Key	improper Please er genetic c selected	eature conta internal stop nsure that th ode table ha or whether q ccept' or 'cod red.	o codon. e correct Is been CDS Jualifiers			Sequence 1		

Figure 107: CDS feature error for improper internal stop codon

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 108, the pencil icon for the 'source' feature is highlighted.

	Seq_3	4	AA	Caballerocotyla klawei	No	
			« <u>1</u> » »			
QUENCE 2						
Sequence Number (ID) 2 Molecule Type DNA Sequence Name ST25_221_BINDING_edited_seq_2 Organism Caballerocotyla biparasitica Length 237				la biparasitica		C
TURES						
ld feature						
id feature eature Key	Location	n		Qualifiers		
	Location	n		Qualifiers protein_id = 3 translation = MYIN		C

Figure 108: Add/edit qualifer Step 1

Clicking on the pencil button highlighted in Figure 108 will open the following feature Overlay, as shown in Figure 109.

Feature 2	
Feature Key source	Feature Location* 1.237
For a detailed explanation of the format of the feature location please $\underline{\operatorname{click}}$ here	
QUALIFIERS Add qualifier 2	
Qualifier Name mol_type	Qualifier Value
Qualifier Name organism	Qualifier Value Cabalierocotyla biparasitica
	Cancel Update feature

Figure 109: Add/edit qualifier Step 2

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 110: the 'Qualifier name'¹⁰ (to be selected from a dropdown list) and the 'Qualifier value'.

¹⁰ See WIPO Standard ST.26 – Annex I, Sections 6 & 8 for all possible Qualifiers for Nucleotides and Amino acids

1	oade	8 ڊ	0
	page	.0	v

QUALIFIERS		
Qualifier Name mol_type	Qualifier Value genomic DNA	Ø
Qualifier Name organism	Qualifier Value Tinamus osgoodi	Ø
Qualifier Name	Qualifier Value	
cell_line cell_type chromosome clone clone_lib	Cancel Create Q	ualifier

Figure 110: Qualifier edit panel

The Qualifier Value field will have a different behavior 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 in Figure 111:

Qualifier Name organelle	*	Qualifier Value *
		chromatophore
		hydrogenosome
		mitochondrion
		nucleomorph

Figure 111: Qualifiers with predefined values

- 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 in Figure 112:

	NOTE: this value may require translation for National/Regional procedures.				
	Non English Qualifier Value				
Ŧ					
	v				

Figure 112: Qualifiers with free text values

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.

- **Qualifiers with pre-defined format**. The value field is a free-text field, but the value entered is validated to ensure it matches the specific rules provided in WIPO ST.26 Annex I, Section6¹¹. In the example shown in Figure 113, the date has not been provided in the correct format:

Qualifier Name collection_date	•	Qualifier Value 2002-08-
		Invalid qualifier value format. The value for the qualifier 'collection_date' must have the format 'YYYY-MM-DD', 'YYYY-MM' or 'YYYY'.

Figure 113: Qualifiers with predefined format

- **Qualifiers with no value allowed**. The qualifier value field is not editable, as indicated below in the example in Figure 114:

Qualifier Name germline	Ŧ	Qualifier Value	
Figure 114: Qu	alifiers	s 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 115 to proceed.

Feature Key Source For a detailed explanation of the format of the feature location please <u>click hera</u>	Feature Location* 1339	
For a detailed explanation of the format of the feature location please <u>click here</u>		
QUALIFIERS Add qualifier Qualifier Name moLtype	Qualifier Value other DNA	Ø
Qualifier Name organism	Qualifier Value synthetic construct	0

Figure 115: Update feature after qualifier edit

¹¹ See WIPO Standard ST.26 – Annex I, Section 6 for rules regarding these qualifiers.

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 Figure 34 .

WIPO Sequence	PROJECT Demo	VERIFICATION REPORT	LANGUAGE DEPENDENT QUALIFIERS	IMPORT Report	DISPLAY THE SEQUENCE LISTING	HELP	•	PREFERENCES	EI₹	Return to project home
GENERAL INFORMATION	SEQUENC	ES								
PROJECT	Г	Print	Export	nport Anoth	er Project	Valid	ate	Generate Seq	uence Lis	sting
Project Name project der Status modified Description Original free text languag Automatically add a trans created On	je code	ier when a CDS fe	F S	ile Name Sequences 2	e 2022-01-12 2 free text langu	lage cod	de			Ø

Figure 116: View verification report Step 1

Depending on whether the project sequence listing is valid or not there will be two resulting Views: Figure 117, where errors/warnings have been generated and Figure 118, after a successful validation.

IPO Sequence User Man	ual					page 83
WIPO Sequence	PROJECT VERIFICATION Demo Report	LANGUAGE IMPORT DEPENDENT REPORT QUALIFIERS	DISPLAY THE SEQUENCE LISTING	HELP 👻	PREFERENCES	El v Return to project home
Report genera	ited on 2022-01	-12				Print Report
Severity ٥	Data Element o	Message Text	Detecte	ed Value	Detected	Sequence
ERROR	Application Identification	The mandatory IP Office Code is missing.	-			
WARNING	Earliest Priority Application Identifications	Priority application information has been entered, but no prior application has been designated as the earliest. The Earliest priority application must be designated when a priority claim is made to an earlier application.	-			

Figure 117: Example verification report, errors/warnings

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 117. The generated report will be displayed in the PDF viewer as shown in Figure 9.

WIPO Sequence TES	ERIFICATION EPORT	DEPENDENT QUALIFIERS	IMPORT REPORT	THE SEQUENCE Listing	HELP▼	PREFERENCES	[-	to project home
Report successfully generat								

Figure 118: Verification report, no errors/warnings

LANGUAGE DEPENDENT QUALIFIERS

This section details the functionalities provided in the Language Dependent Qualifiers View. 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 119.

WIPO Se	quence TRANSL	VERIFICATION REPORT	LANGUAGE DEPENDENT QUALIFIERS	IMPORT REPORT	DISPLAY THE SEQUENCE LISTING	HELP 👻	PREFERENCES	E▼ Return to project home
				IMPOR	T FREE TEXT	QUALIFIERS	EXPORT FREE T	EXT QUALIFIERS
LANG	UAGE	DEPEN	NDEN	IT C	UAL	IFIE	RS	
Source langua	ge code for free text qu	Jalifiers en	Т	arget langu	lage code for f	ree text qual	fiers	Ø
Sequence ID Number	Sequence Name	Feature Key	Feature Locatio		lualifier)	Qualifier Name	Qualifier Value	Non English Qualifier Value-(e
1	transl_except_se q_1	source	164	q	1	organism	Sialia currucoide	

Figure 119: 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 110, 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.

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 120: Example of 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 121 to Figure 125).

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

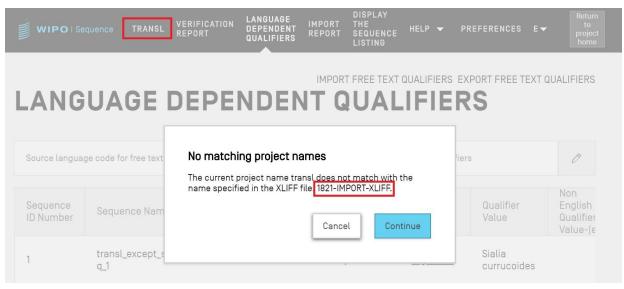


Figure 121: Project name validation

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

WIPO Sequence	User Manual			page 86
ce	Somo qualifier IDa oc	uld not be menned		
	Some qualifier IDs co The following qualifier IDs		napped to those in the current projec	t.
	Qualifier ID	Source	Target	
	q2	Una virus	organism fr	
		₩ 4 1	₩	
			Cancel	tinue

Figure 122: Qualifiers mapping validation step

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

Feature Key	Detected target language code The detected target language code is "en". The free text translations provided will be imported and update the	Value
source	existing qualifier values with the qualifier values indicated in the input file.	omyces ie
source	Cancel Proceed	piens

Figure 123:Source language validation step

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

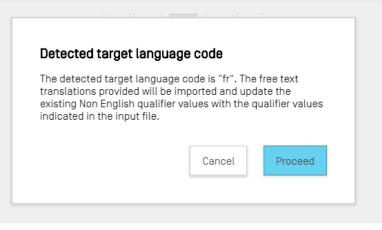


Figure 124: Target language validation step

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.

WIPO	quence TRAI	NSL VERIFICATI Report	ON LANGUAGE DEPENDENT QUALIFIERS	IMPORT REPORT	DISPLAY THE SEQUENCE LISTING	HELP 🔻	PREFERENCES	E ▼ Return to project home
SUCCESS: The	e free text qua	lifiers have bee	en imported cor	rrectly				×
IMPO	RT FF	REETE	EXT QI	JALI	FIER		EPOR	
Sequence ID Number	Feature Key	Feature Location	Qualifier ID	Qualifier Name	Origina qualifie value		English	Imported qualifier value
2	source	119	q3	organism	Sialia curruce	oides		Construction synthétique
1	source	164	q1	organism	Sialia curruce	oides		
4			M M	1	₩			*

Figure 125: 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 125.

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 values¹².
- The associated Qualifier and Feature information provided in the Table shown in Figure 119.

This file can be viewed, edited and imported in the tool again after providing the appropriate translation following the steps shown in Figure 121 - 125.

IMPORT REPORT

This section details the functionalities provided in the Import Report View.

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 126.

	REPORT	QUALIFIERS	REPORT	SEQUENCE Listing	HELP		PREFERENCES	ENT	project home
The import report has not bee A sequence listing can be imp	• •	nport Sequence	Listing" but	ton in the proje	ect list ho	ome p:	age.		

Figure 126: 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 Figure 127)
- Changed Data Table (see Figure 128)

¹² The translation of the selected Non English free text language code provided in the project during export

Import Report Table

WIPO Sequence	2469- RNA_ST25	VERIFICATION REPORT	LANGUAGE DEPENDENT QUALIFIERS	IMPORT REPORT	DISPLAY THE SEQUENCE LISTING	HELP	•	PREFERENCES	E▼	Return to project home
Import Re	port								Print I	Report
Import Report	Messa	iges								
Type of Note	Data E	lement	Messa	ige Text			Dete	cted Sequence		
Global	<130>		File Re autom follow	eference' h natically as ing value: can be edi	ne 'Applicant has been ssigned the USPTO; this ted within th					
Global	<110>		must l applic be sel togeth	oe selecte ant or one ected as t	n or person d as the first person mus ne inventor, e appropriate	t				

Figure 127: Example import report

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

- **Type of note**: "INDIVIDUAL" for a message related to a specific sequence <u>or</u> "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).

Changed Data Table

Global	<	400>	been re	etter amino acid symb placed with their corre ter codes.			
				2 🇯 🗰			
Changed	Data						
Origin Tag	Origin Element Name	Origin Element Value	Target Element Name	Target Element Value	Transformation	Origin Sequence ID	Sequence ID Number
<221>	Name/Key		Feature Key	misc_feature	The custom feature key has been replaced with a recommended key (see Annex VII, ST.26).	1	1
<223>	Other information		Qualifier Name	note	A 'note' Qualifier has been created.	1	1

Figure 128: Example Changed 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 127):

- 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.

DISPLAY SEQUENCE LISTING

This section details the functionality available in the Display the Sequence Listing View.

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 129).

April 2023

For further details on how to generate a sequence listing, go to Figure 63.

WIPO Sequence	TESTSTRAIN	VERIFICATION REPORT	LANGUAGE DEPENDENT QUALIFIERS	IMPORT Report	DISPLAY THE SEQUENCE LISTING	HELP▼	PREFERENCES	[▼	Return to project home
ST.26 sequence listing				able format					
				Export	Sequence List	ing as .txt fi	ile Display S	equenc	e Listing

Figure 129: Display sequence listing, generated successfully

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 130.

WIPO Sequence	PROJECT Demo	VERIFICATION REPORT	LANGUAGE DEPENDENT QUALIFIERS	IMPORT REPORT	DISPLAY THE SEQUENCE LISTING	HELP	•	PREFERENCES	EI▼	Return to project home
File not found.	displayed in I	human readable f		quence listi	ng has been g	enerate	d in XI	4L format. Please	return to t	ha
project page and sele		Sequence Listing'.								

Figure 130: Display sequence listing, file not found

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 131.

Note:

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

C:\Users	AppData\Roaming\ST26_authoring\Project_6\generated.html	- ℃ Search_
-		
ence	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

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 132: Display sequence listing, Example TXT
```

If the generated sequence listing, in XML format, is greater than 100Mb in size, instead of

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 133 will be displayed.

~ -

Generated sequence listing × +

← → C ① © File | C:/Users/aterrass/AppData/Roaming/ST26_authoring/Project_1065/generated.htm

ST26 SEQUENCE LISTING

The xml generated is bigger than 100 MB therefore it is not possible to display the sequence listing in a human readable form.

Figure 133: Display sequence listing, HTML too large to display

5 FILE FORMATS

The following file formats can be imported into WIPO Sequence.

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

MULTI-SEQUENCE

The multi-sequence format¹³ 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>

uuuucuuauuguuucuccuacugcuuaucauaaugauugucguaguggcuuccucaucgucuccccacc gccuaccaacgacugccgcagcggauuacuaauaguaucaccaacagcauaacaaaaagaaugacgaa gaggguugcugauggugucgccgacggcguagcagaaggaguggcggagggg

<Second Sequence; DNA; Albies alba>

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.

¹³ <u>https://www.uspto.gov/sites/default/files/patents/resources/tools/checker/patentin351_20110214__6_.pdf</u>

Example:

FASTA

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

FASTA file with one sequence

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

CTCAACAATCTGAAGCATCG

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

[End of Document]