Omixon Target User Manual

Omixon Target User Manual

Introduction

Introduction

Omixon Target ('the tool') is a suite of software for analysing targeted sequencing data from next generation sequencing (NGS) platforms.

The main vision behind this tool is to help laboratories move towards using NGS data for the analysis of diagnostic targets.

There is a Quick Start Guide for Omixon Target available on the Omixon website (http://www.omixon.com/web/guest/omixon/targetquickstart.htm).

Omixon Target is for Research Use Only. Not for use in diagnostic procedures.

Uses

Analysis of Diagnostic Targets

HLA Typing

Omixon Target has a simple HLA Typer for NGS data. This tool is easy to set up and run, and multiple samples can be analysed together. It provides both a visual summary, and tables of reports with statistical confidence measures for the accuracy of the results.

Data Analysis

Omixon Target includes preconfigured support for analysing some of the most common human diagnostic targets such as HLA, CFTR, and BRCA, however you can also configure the tool to analyse any gene or region of interest. The underlying mapping, alignment and variant calling algorithms are intended to meet the high precision and analysis quality control requirements of diagnostics labs. You can easily move from a high-level mutation summary table to inspect the underlying short read data, and based on this visual inspection you can manually approve variants. The tool helps to identify amplification artefacts and other error sources using a simulation-based validation track as an analysis control. The tool can be used for multi-sample or family trio comparative studies.

Discovery

Omixon also use the tool to support collaborative genomic biomarker discovery projects.

Sequencing Technologies

Omixon Target supports sequencing data from these major sequencing technologies:

- Illumina
- Ion Torrent
- Roche 454

Variant Calling

The only Variant Caller that is available in the Standard edition is a GATK pipeline (using the Broad Institute GATK variant caller). This gives excellent results with Illumina data. There is a better variant calling pipeline for Ion Torrent or Roche 454 data available within the Pro edition.

HLA Typing

So far, most of the validation that has been done on the prototype HLA Typing algorithms has been done with paired Illumina data. While it is possible to get results with other kinds of data, this is not yet reached an equivalent standard as using the paired Illumina data. The HLA Typing results for the other platforms will be improved in future releases.

Modules

Omixon Target has a modular structure

Omixon Target consists of a main Genome Browser Module, with some optional Modules.

This means that Omixon Target can simply be used as a Genome Browser, if desired (i.e. for visualization of NGS data analysis results, in SAM/BAM and VCF formats).

Genome Browser Module

This is always included as standard with Omixon Target. It includes some data management functions, as well as a full-featured NGS Genome Browser and some tables for variant visualiziation.

The Genome Browser Module works on Windows, Linux and Mac OS X.

Standard Data Analysis Module (Optional)

This module includes a range of general genomic data analysis algorithms, including Omixon's own aligner and a full GATK variant call pipeline. It also features an approval work flow, for automated or manual filtering and approval of variants.

The Standard Module works on Windows (64 bit recommended), Linux (64 bit recommended) and Mac OS X.

Pro Data Analysis Module (Optional)

This includes all the features of the Standard Data Analysis Module, including all the standard genomics data analysis algorithms and the approval work flow. This Module is an alternative to the Standard Data Analysis Module, for Linux users.

In addition, it also includes integrated access to a number of third party tools, including a better variant call pipeline for Ion Torrent and Roche 454 data (based on the SamTools mpileup and some supporting scripts), a demultiplexing import and an sff import feature. The BWA aligner for Illumina data and a simulation-based Quality Control feature will be coming soon.

The Pro Module is only available for a 64 bit Linux operating system and requires some manual installation and configuration of third party tools in addition to the usual installation steps.

HLA Typing Module (Optional)

This Module contains only the HLA Typing feature of Omixon Target (i.e. it has none of the data analysis features).

The HLA edition has no up-front license fee. The HLA Typing Module offers two licensing schemes:

1. A credit-based pricing scheme with no up-front license fee, with limited use (one credit is consumed per sample analysed, more credits can be purchased on demand). 40 free HLA typing credits are included in the evaluation version.
2. An annual fee based license, which allow unlimited HLA Typing. Please contact sales@omixon.com for a quote.

The HLA Module works on Windows (64 bit recommended), Linux (64 bit recommended) and Mac OS X.

Adding and removing Modules

Note that it is possible to add and remove Modules to and from your Omixon Target installation. The Modules are controlled by a license file, so if you would like to alter your installation then you should contact Omixon and we can generate a new license file for you, which can be imported within the Settings function in the tool.

Key Concepts

Target

The data analysis part of the tool is built around the concept of an analysis target. This can currently be one or more genes, exons or amplicons.

There is a Target Dashboard to manage the targets within the tool. The annotations for a Target can be imported directly from a bed or gff file.

While using a target is recommended it is not actually mandatory within the tool.

Hierarchy of Experiment, Analysis and Sample

In order to help with organising the data belonging to the samples that are sequenced, there is a simple data management hierarchy in the tool.

The sequencing data will belong to a 'Sample'. A number of samples can be grouped together within a single 'Analysis'. And finally, a number of analyses can be grouped together within a single 'Experiment'.

The configuration of shared data, such as:

- the target regions for the investigation
- which reference species genomes will be used (and the imported reference sequence data and known variations)
- and the sequencer used for the sequencing of the data

can be done at both Experiment and Analysis level.

If desired, a more generic configuration can be created within an Experiment, with a more specific configuration chosen within each Analysis. For example, it might be desirable to separate the Analyses by sequencer, if multiple sequencers are being used. It's also possible to use a sub-set of targets within the Analysis, for example the HLA-A gene could be the target of one Analysis, and the HLA-B gene could be the target of another.

There is an Experiment Dashboard, an Analysis Dashboard and a Sample Dashboard to manage each of these.

Profiles

The easiest way to use Omixon Target for data analysis is to use one of the pre-built, pre-configured Profiles. This already contains a Target definition and all the reference data required for a particular Target. Adding a new profile will cause the data to be downloaded and automatically configured with the tool. It will also create and configure a starting Experiment, Analysis and Sample, plus import some example Sample data if required.

Reference Genomes

One or more reference genomes can be configured for each Experiment. One or more 'contigs' (usually chromosomes) can be chosen for each reference genome. The reference genomes and/or contigs can split across the Analyses.

The sequence data for the reference genome (fasta format) will need to be imported into the tool. For the best variant analysis results, the dbSNP (known mutation) data for the reference (if available) can also be imported into the tool.

There is a Reference Genome Dashboard to manage the data for the reference genomes.

Samples

Each 'Sample' will generally consist of the sequencing data or results for a single individual. There are three possibilities for importing sample data:

- The 'raw' sequencing data (short reads) can be imported into the tool (in fastq format), after which it can be mapped and aligned against one of the reference genomes (one or more contigs) and variants can be 'called' (detected) - all in a single step.
- It is also possible to import short reads that have been mapped by another tool and just do the variant detection - or simply just browse the data in the Genome Browser.
- Finally, variant call results can be imported alone, without any short read data at all, and visualised and analysed within the tool.

Once variants have been called, there are a number of downstream analysis steps that can be performed, including browsing the results in the Genome Browser, analysing the results within the Analyse Genome Dashboard, 'approval' of variants via the Approval Dashboard, and creating and viewing reports in the Report Dashboard.

Mapping, Alignment, Variant Call

These functions are only available within one of the optional Data Analysis Modules (Standard or Pro).

One of the primary goals of the tool is to identify variants within the sample.

The tool includes sequencer-specific algorithms for mapping and aligning NGS short reads against a reference sequence, and then calling variants against the aligned short reads using a GATK pipeline (following the recommended best practises of the Broad Institute).

The mapping and alignment algorithms used are currently Omixon's own. There are separate algorithms for Illumina, 454 and Ion Torrent data, which include different error models for the different sequencers. One future plan is to run not just Omixon's algorithms, but some other alignment algorithms as well, and to compare and/or merge the results of this double alignment.

It is also possible to simply import already mapped short read data and just run a variant call, or even just import already called variants into the tool, for downstream analysis.

Genome Browser

There is a genome browser included in the tool. This allows you to browse the selected target regions, manually inspect the variants and short reads, see a consensus sequence, and see the coverage of the short reads. Multiple samples can be browsed together.

Analyse Sample Variants

It is also possible to import 'expected' variants into the tool. There is a function called 'Analyse Sample' which will automatically compare the expected variants with the actual variants found (either imported or via the map/align/call variants functions). These expected variants can come from previous analyses (for example with Sanger data), or they are also useful if the data being analysed is simulated data, with known results.

The Analyse Sample Variants function also allows the results of the Quality Control to be inspected.

Sample Difference

If multiple samples are selected in the Analysis Dashboard they can be compared against each other using the 'Compare Samples' function, which opens the Sample Difference screen. They can also be browsed together and compared in the Genome Browser.

Manual Approval

This function is only available within one of the optional Data Analysis Modules (Standard or Pro).

It's possible to 'approve' or 'reject' variations within a Sample. Approved variations will become candidates for downstream analysis, such as HLA typing. Rejected variants will be ignored by downstream tools. Approval and rejection can be done either by the Automatic Approval or Automatic Rejection functions, or via a manual approval process within the Approval Dashboard screen.

Bioinformatics Notes

Reference Sequences

By default, Omixon Target is set up to use HG19 as the main human reference sequence. It is however possible to configure the tool to use another reference sequence.

Alignment Tools

These tools are only available in the optional Data Analysis Modules (Standard and Pro).

The main underlying tool for the 'Map and Align' step is the Omixon Variant Toolkit. This is also available as a standalone, command line tool. The Toolkit uses a 'properties' file for it's parameters, and the same file can be used to transfer 'advanced' parameters to the underlying Toolkit while running the 'Map and Align' step within Omixon Target. The Toolkit has it's own readme file, which is also available for download via the toolkit page on the Omixon web site: https://www.omixon.com/omixon/abouttoolkit.htm

BWA will be added soon to the Pro Module.

Variant Calling Tools

These tools are only available in the optional Data Analysis Modules (Standard and Pro).

Omixon currently use open source, third party variant calling tools. The only one included in the Standard Module is the GATK variant caller from the Broad Institute. Also included in the Pro Module is the Samtools variant caller (mpileup), which gives better results for Ion Torrent and 454 data.

System requirements

Omixon Target is supported on 3 platforms:

- Mac OS X
- Linux
- Windows

The recommended hardware requirements for all the optional modules are the following:

- 64 bit multi-core CPU with 64 bit operating system,
- at least 8GB memory,
- storage space requirements mostly depend on the size of data usually used for analyses.

The memory requirements depend on the size of the reference genomes being used.

The memory limitation of 32 bit operating systems is the primary restriction for not fully supporting these. For the Genome Browser Module only or for demonstration purposes with small genomes, the tool can still be used on a 32 bit operating system with up to 4GB memory.

An Oracle Java 6 Runtime Environment (JRE) is required, this is always included inside the installer.

Installation guide

There are separate installation packages for all three supported operating systems.

In the following sections are descriptions for the installation steps required for each operating system.

The Omixon Target Pro version of the product has also a separate section to summarize installation and setup instructions related to that specific type of distribution.

Windows installation guide

There are many versions of Windows operating systems. We have tested Omixon Target with Windows 7, Windows XP and Windows 8.

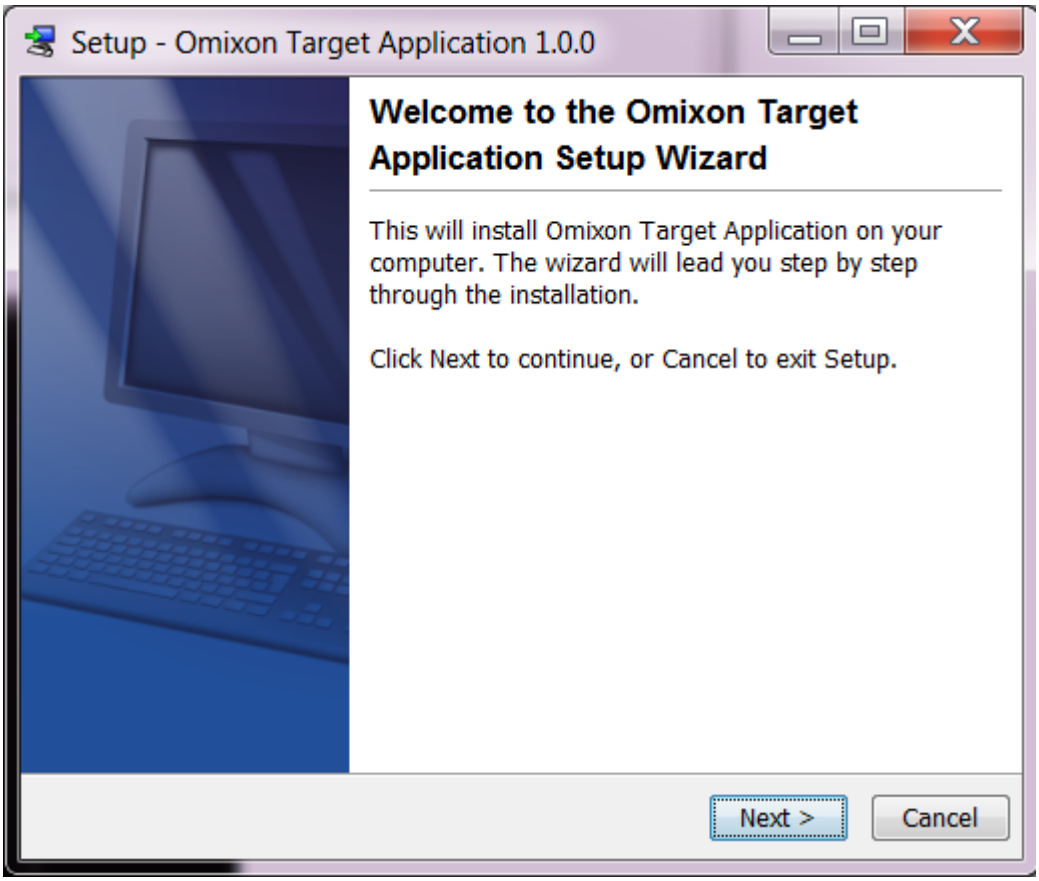
We provide two versions of the installer package for Windows system, both are bundled with a Java Runtime environment:

- a 64 bit version, named omixon-target_windows-x64_1_0_0_with_jre.exe,
- a 32 bit version, named omixon-target_windows_x86_1_0_0_with_jre.exe.

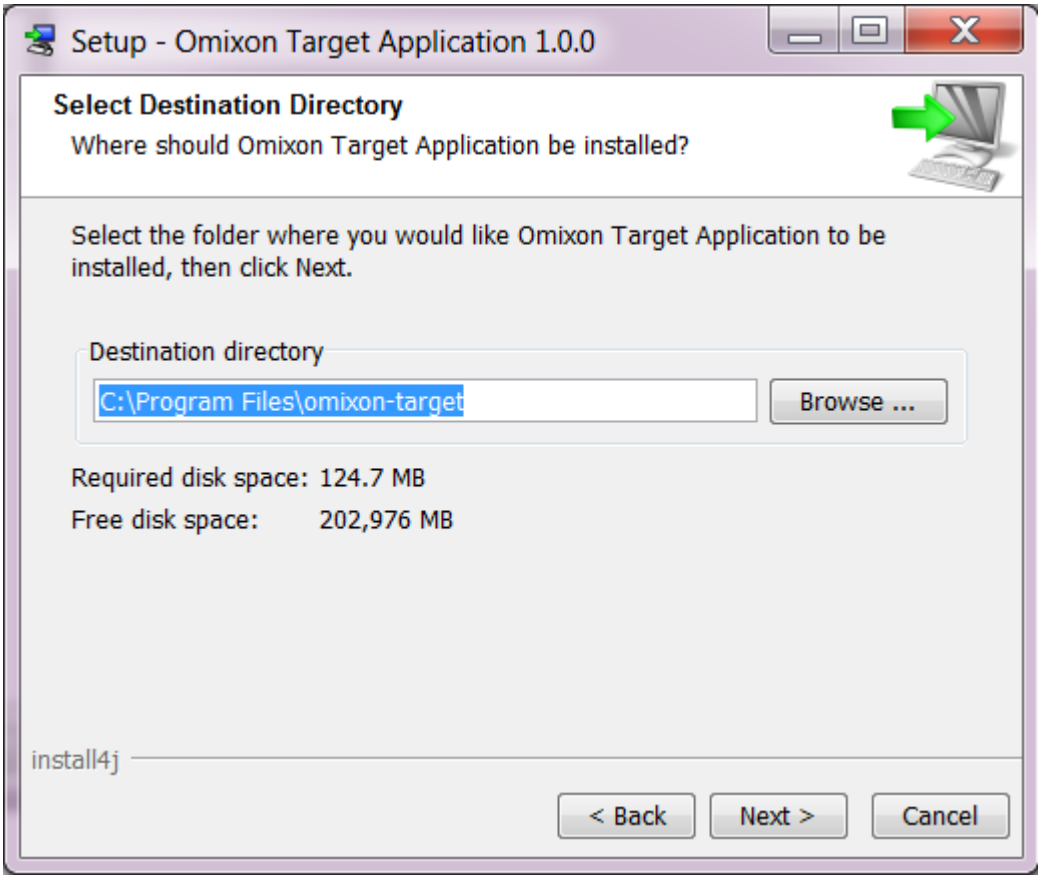
The only fully tested operating systems are the 64 bit ones.

The installer package which contains Java Runtime environment contains a 64 bit JRE and it is recommended for 64 bit operating systems (omixon-target_windows-x64_1_0_0.exe).

After running the installer, the following dialog should appear.

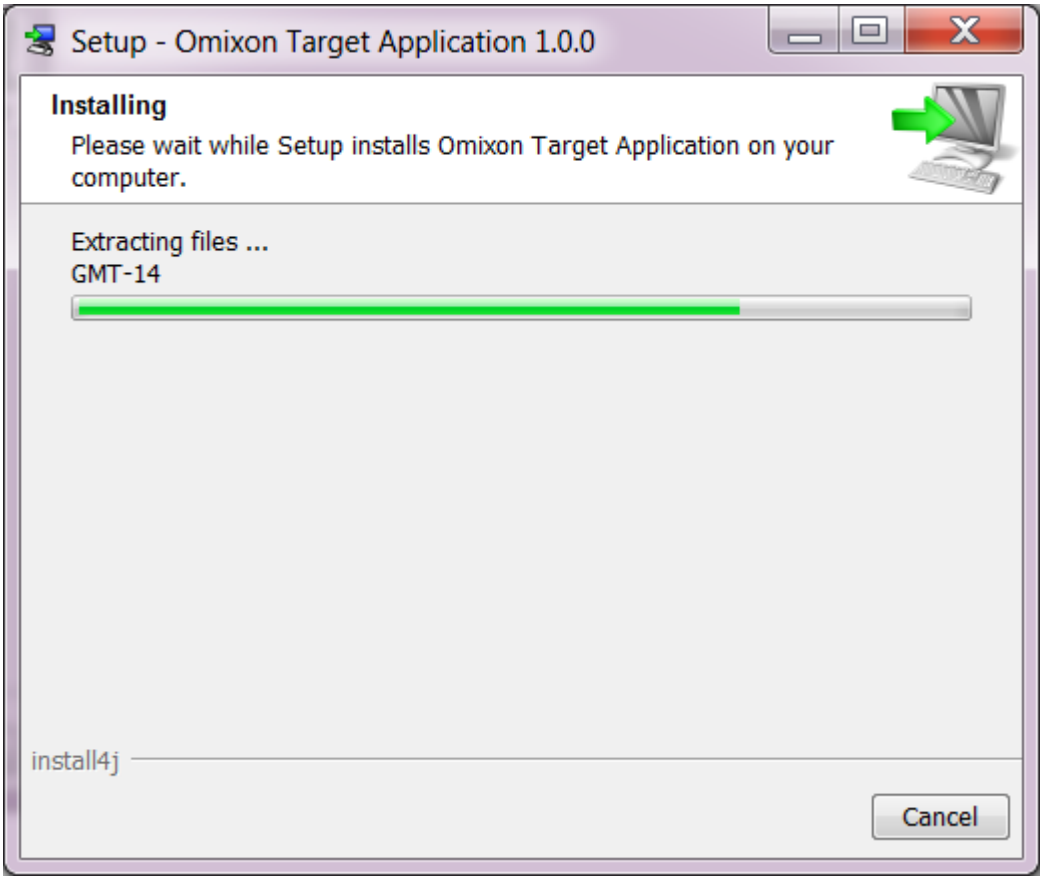


The next step is the destination directory selection.

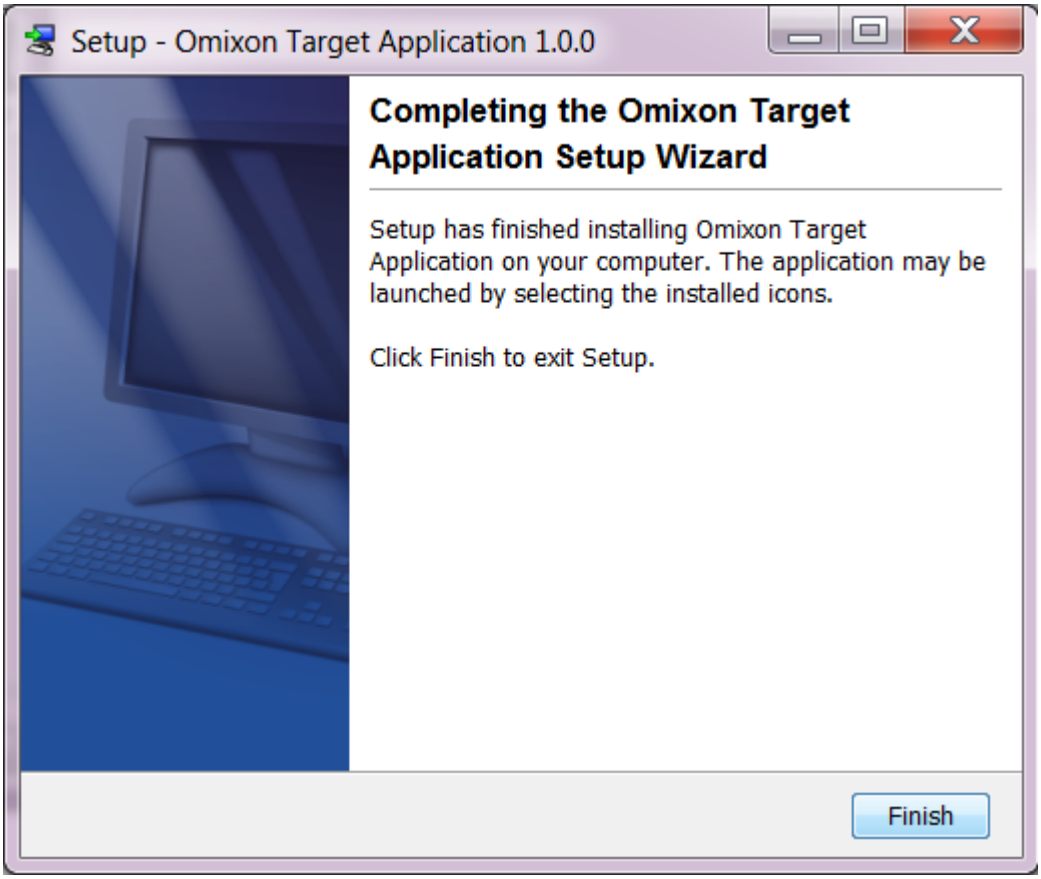


By default, the Program Files directory is selected, in addition to that, an application directory is appended.

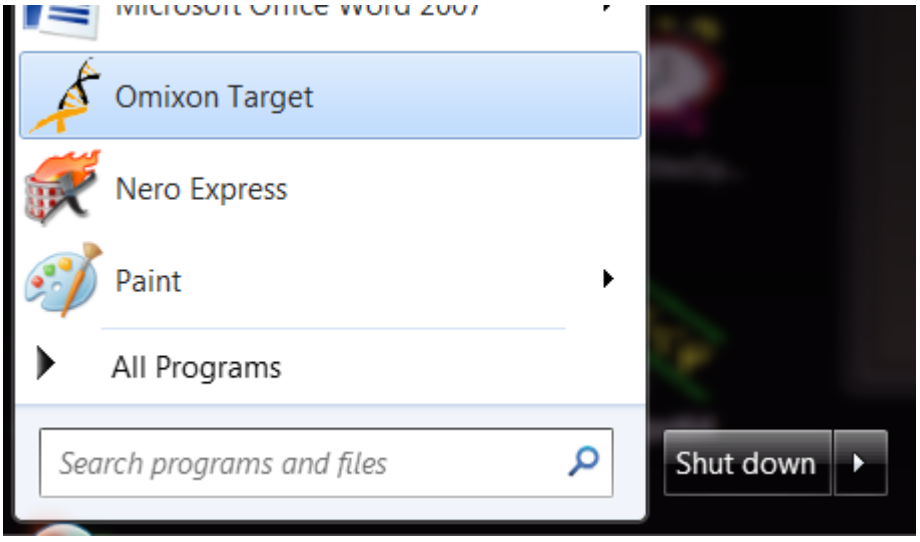
After selection, by clicking on the Next button, the installation process is followed.



At the end of the installation process, the following dialog indicates the success of the installation.



After the installation, the following icon becomes available in the Start menu.



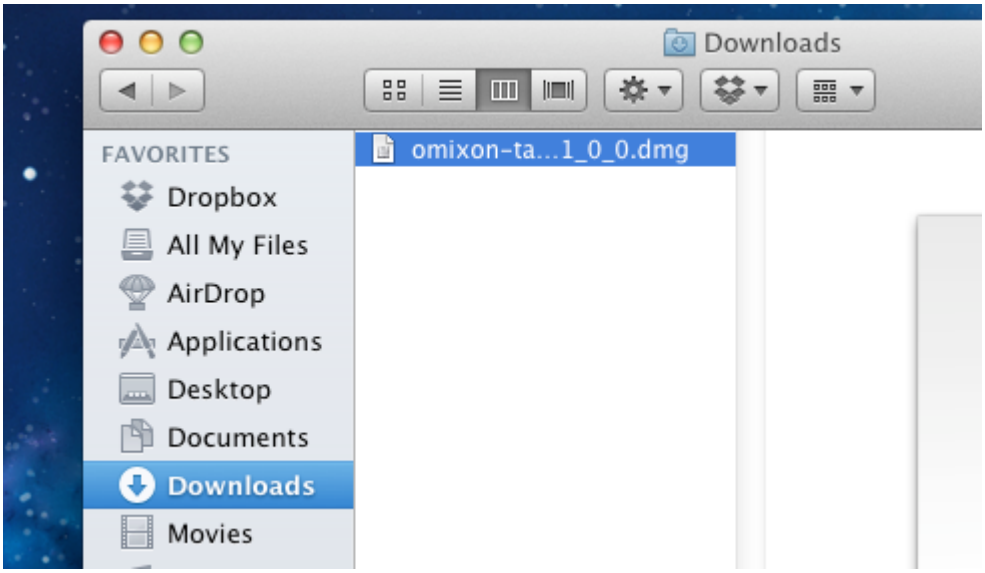
The Omixon Target application can be started by clicking on this startup icon.

Mac OS X installation guide

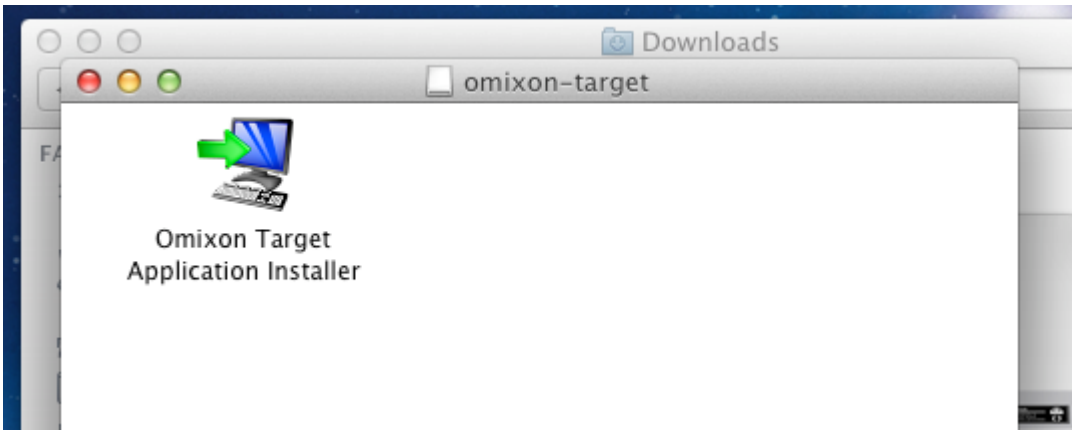
As Mac OS X is an exclusively 64 bit operating system and it always has a proper Java Runtime environment for our tool, therefore the installer does not contains Java Runtime Environment.

The installer is in DMG format.

In "Finder" file manager, the installer file should look like this:

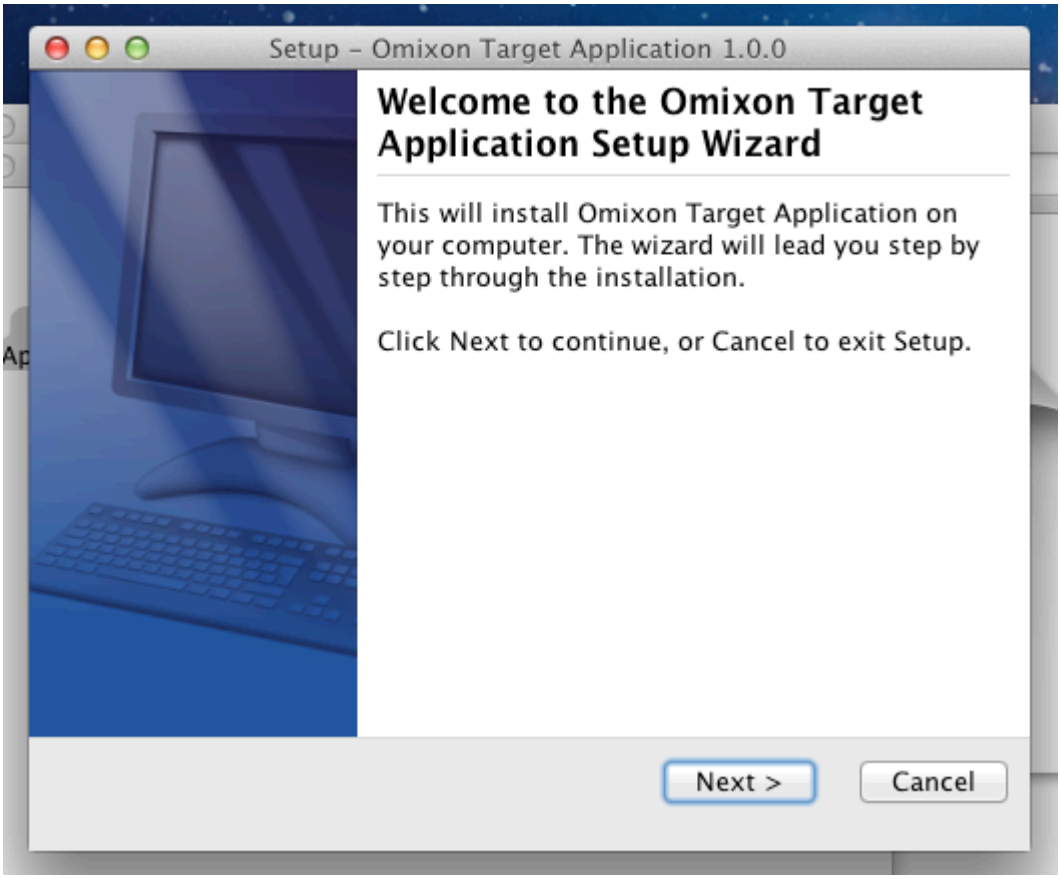


Once the DMG archive is opened, you can see the installer icon.

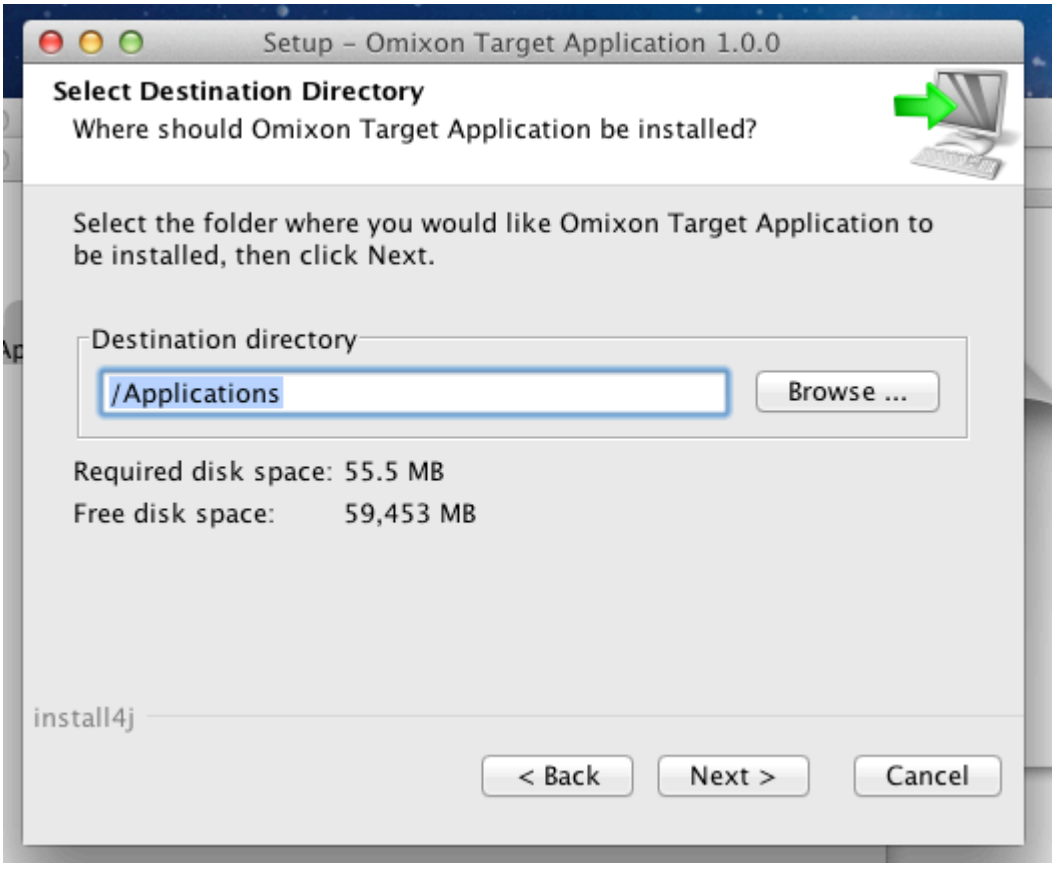


This has to be started, like any other Mac OS X installers.

Note that if you **don't** already have Java installed on your Mac, it will automatically be downloaded and installed at this point.

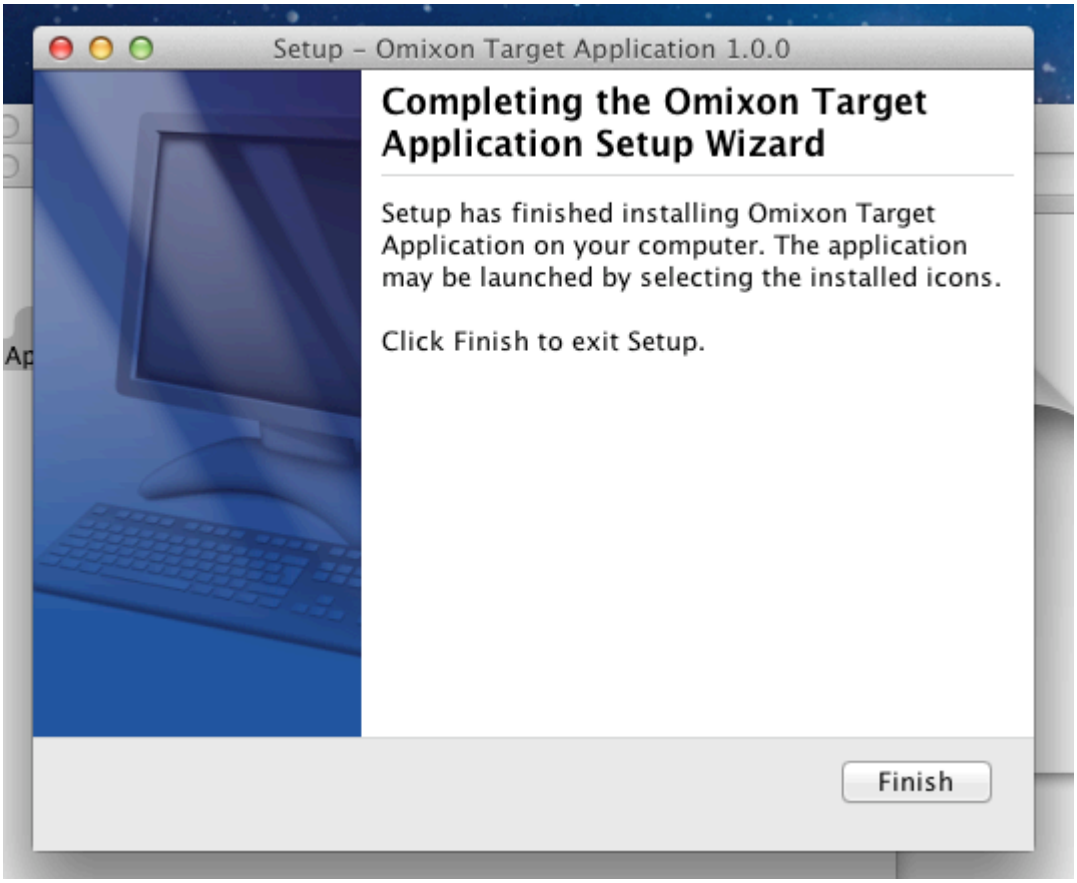


The next step is to set the destination directory.

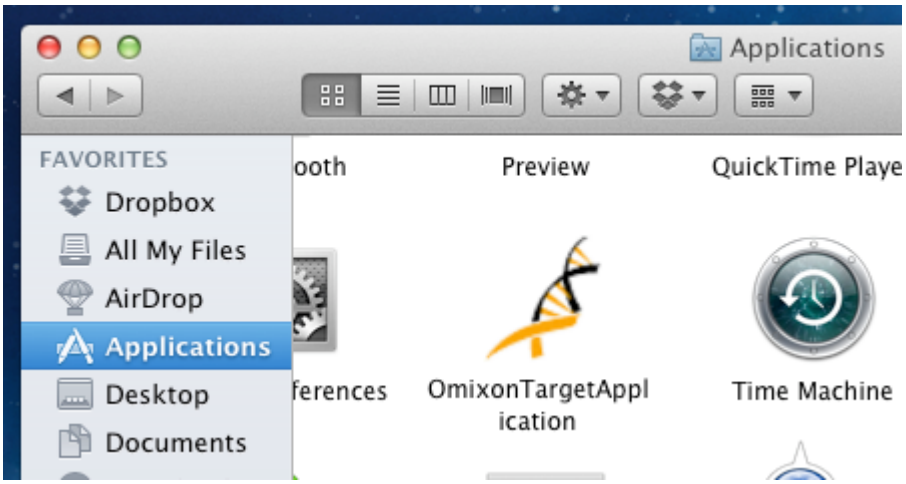


The default destination directory on Mac OS X is /Applications.

By clicking on the Next button, the installer performs the installation process and the final dialog will look like the following.



After installation, a startup icon named "OmixonTargetApplication" can be found in the Applications folder.



In Mac OS X, the uninstallation process is straightforward, the user has to just send to Trash by the application icon.

Linux installation guide

We provide two versions of installer package for Linux operating systems:

- 64 bit, with Java Runtime environment,
- 32 bit, with Java Runtime environment.

The installer package which contains the Java Runtime environment contains a 64 bit JRE and it is recommended for 64 bit operating systems.

In case Omixon Target really needs to be run in a 32 bit environment, it can be done by choosing the appropriate installer. Note, that 32 bit operating systems are not officially supported (not fully tested), but that does not mean that the application cannot run. The current limitation of 32 bit operating systems is mostly related to the amount of memory available there.

The install packages are single file shell scripts. This type of installers are the most suitable to install on various Linux distributions.

After you get the installer file, the shell script still does not have the permissions to run directly. You need to open a terminal window to

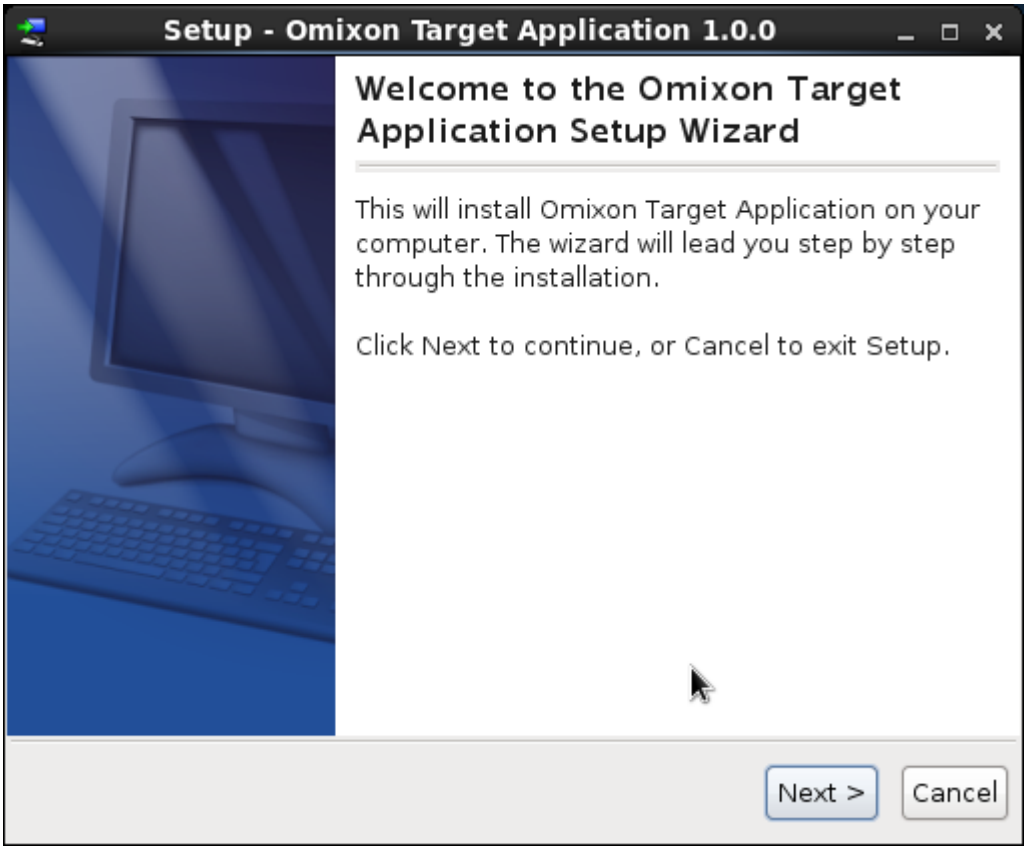
make the installer executable with the following command:

```
chmod +x omixon-target_unix_1_0_0_with_jre.sh
```

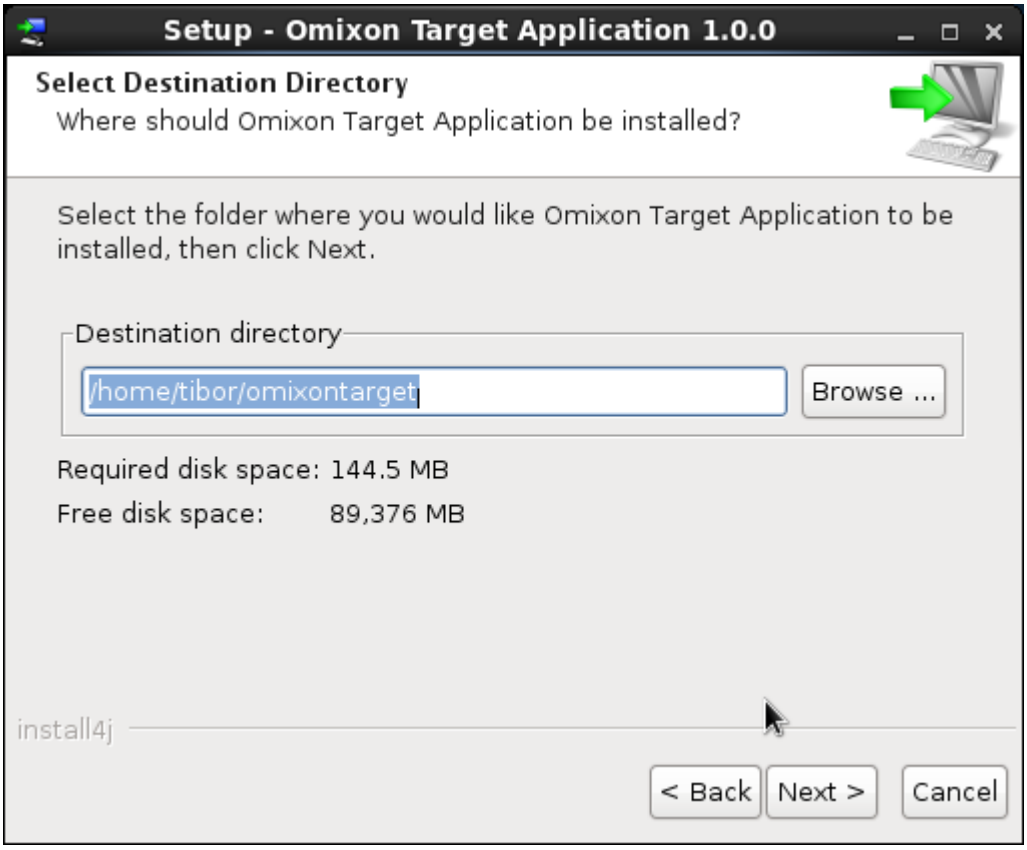
Immediately after that, the installer can be started with the following command:

```
./omixon-target_unix_1_0_0_with_jre.sh
```

Once the installer is started, the following screen appears.

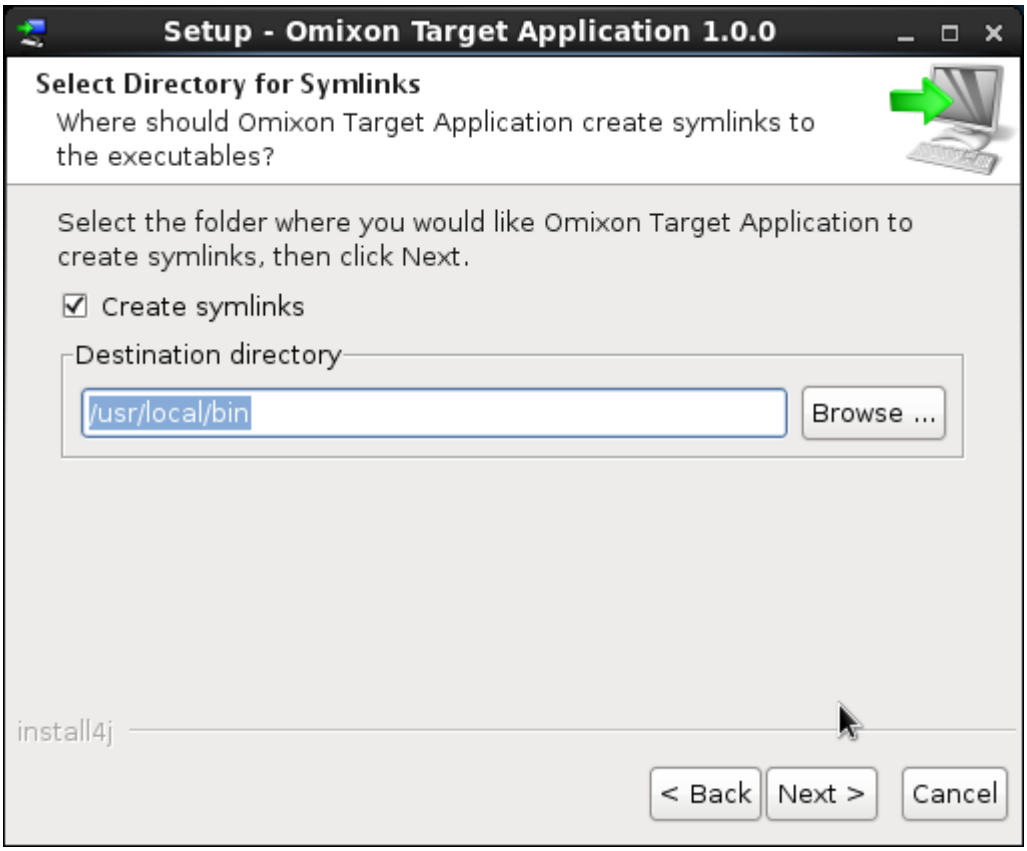


By clicking on Next button, the next step will be the destination directory selection.



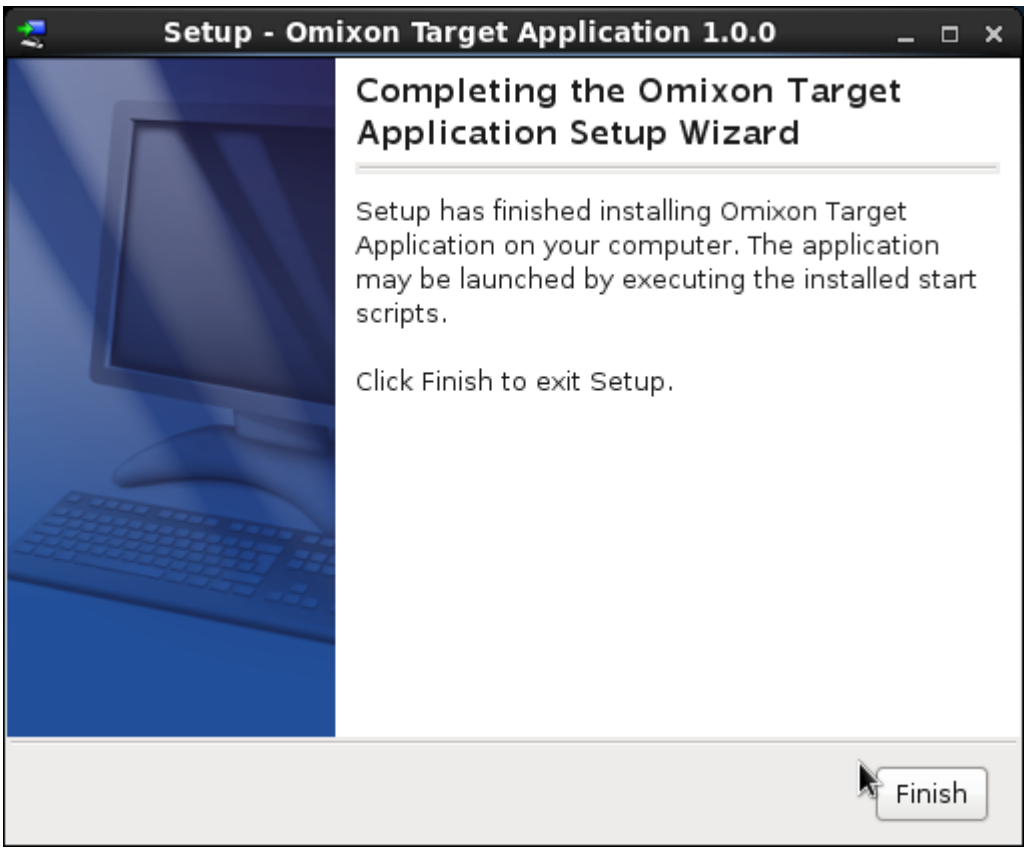
The default destination directory is /home/user/omixontarget. This is the most simple approach to install in various linux distributions.

The next step is symlink location selection.

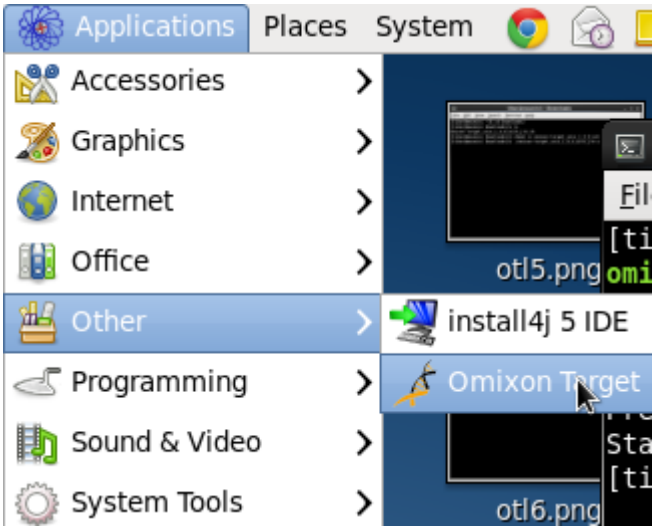


The default location for symlink is /usr/local/bin, but one may choose any other directory. This symlink is useful for users who usually start the application from the command line. Since the installer is creating a startup icon in the system menu, this symlink is optional for most users.

After clicking the Next button, the installation process starts. After it is finished, the final screen appears.



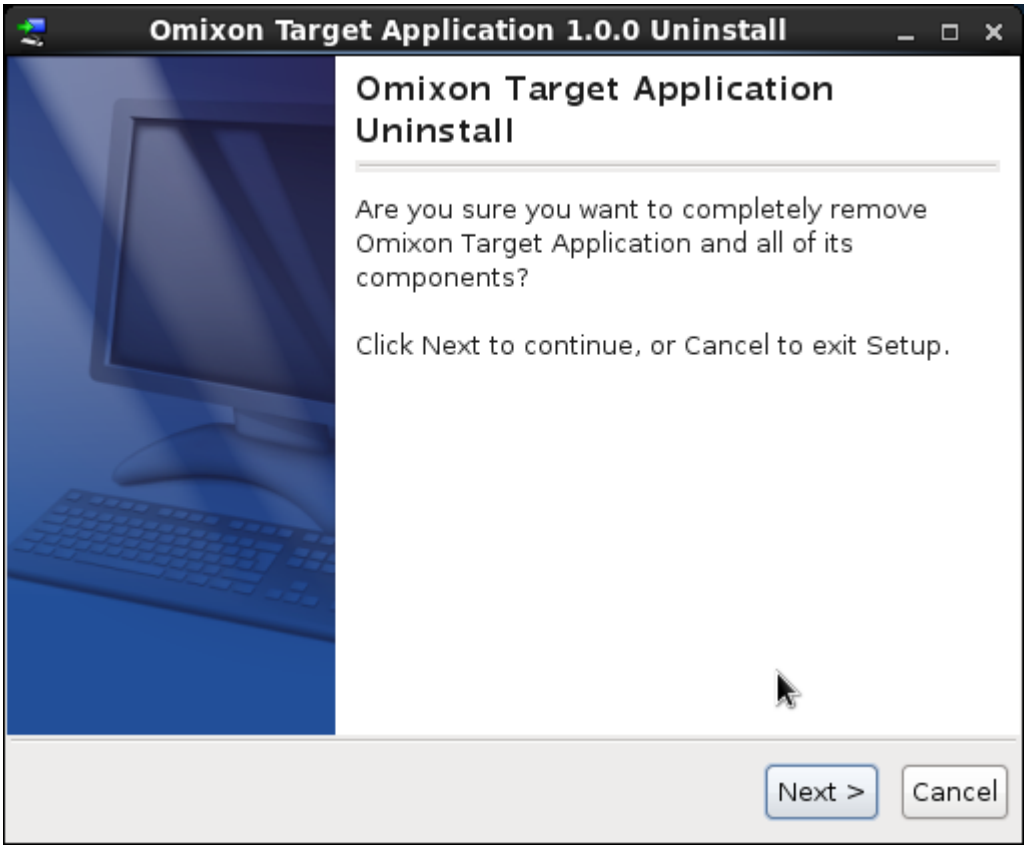
Once the application is installed, an application icon can be found in the System menu.



By clicking on this icon, the Omixon Target application starts and is ready for use.

If you want to uninstall Omixon Target, an uninstaller is available in the directory where the application was originally installed.

The uninstaller shows the following two dialogs.

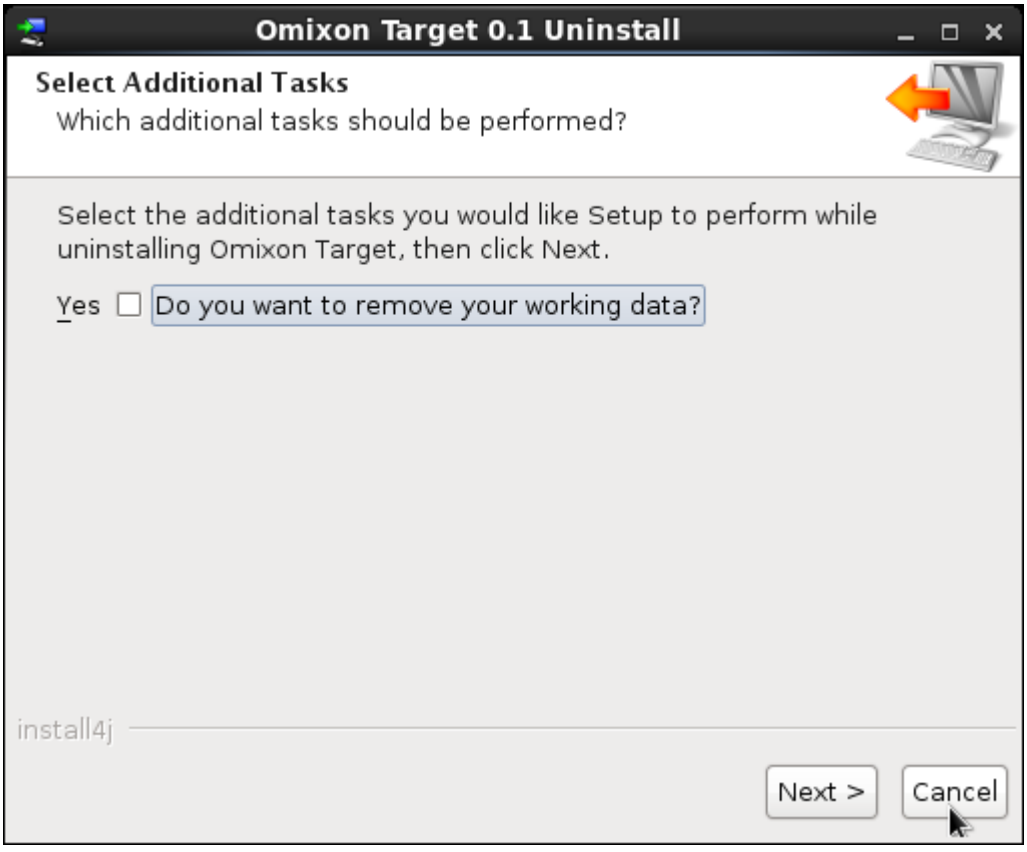


In the next step we are asked about "Do you want to remove your working data?"

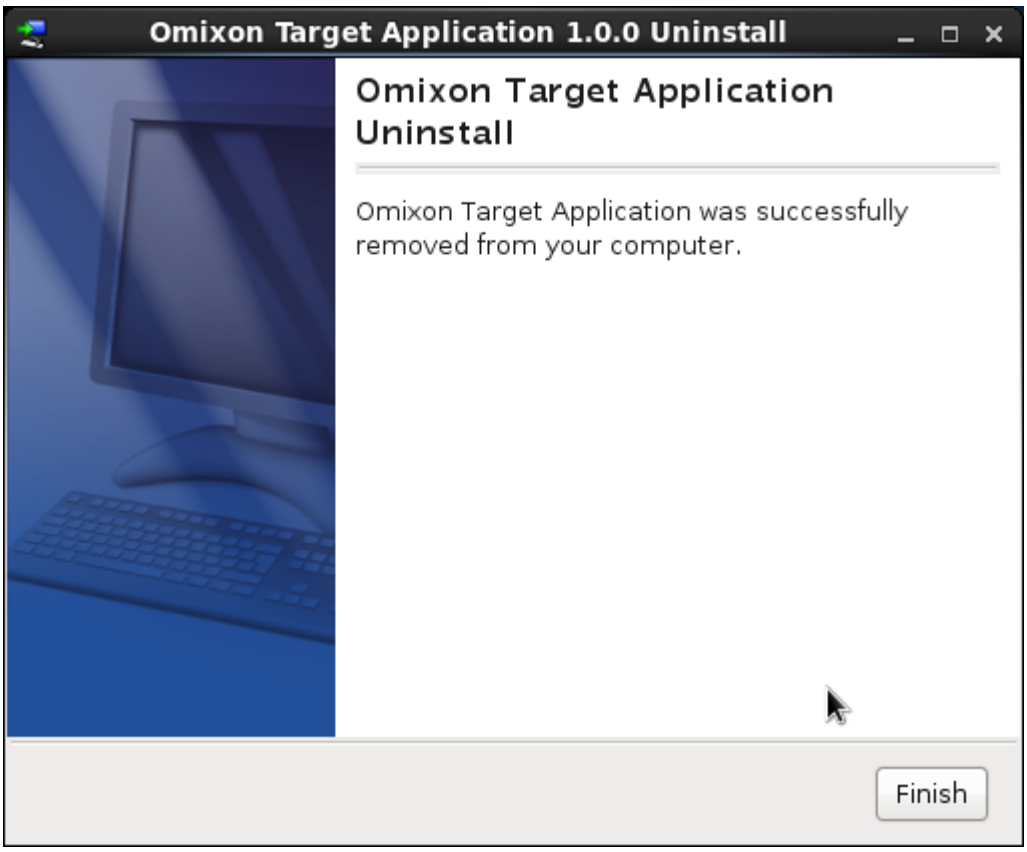
By default the checkbox is not selected, and the directory ~/.ogve/ which is the data directory for Omixon Target, will not be deleted.

If you intend to use future version of Omixon Target and you don't want to be deleted all your work visible in the Omixon Target, then just skip this step.

Instead, if you want to cleanup the working directory, then selet the option 'Yes'.



Finally, the installer informs you about the outcome of uninstall.



Pro Data Analysis Module installation notes

Introduction

The Pro Data Analysis Module is an optional Module which supports additional features compared to the Standard Data Analysis Module.

The Pro Module is supported only on a 64 bit Linux platform.

The Pro Module uses several well-known bioinformatics tools behind the scenes. You need to install and configure these tools manually, once they are on the PATH they can be used by Omixon Target.

General notes and configuration

If you have all the tools installed and available on your PATH then no additional configuration is needed.

If you don't have some tools on your PATH you have to configure them manually for Omixon Target to let the application find them. The configuration file is called **tools.properties** and located in the **proserver/etc** folder of your installation directory. There you can find a configuration line for each tool and you can provide the full path to the executable of the tools.

List of tools

Here is the list of tools to be installed in your environment to get all Pro features working.

Samtools

Homepage of the tool: http://samtools.sourceforge.net/

Version used for testing: *0.1.8*

Installation with package manager on Debian based systems: *apt-get install samtools*

Bowtie

Homepage of the tool: http://bowtie-bio.sourceforge.net/index.shtml

Version used for testing: *0.12.8*

Installation with package manager on Debian based systems: *apt-get install bowtie*

FASTX-Toolkit

Homepage of the tool: http://hannonlab.cshl.edu/fastx_toolkit/

Version used for testing: *0.0.13.2*

Installation with package manager on Debian based systems: *apt-get install fastx-toolkit*

Seqtk

Homepage of the tool: https://github.com/lh3/seqtk

Revision used for testing: *771d60b8f774482a77d60dd8559d17bd487c56e8*

Installation from source:

> *git clone https://github.com/lh3/seqtk.git*
>
> *cd seqtk*
>
> *make*
>
> *cp -f seqtk /usr/local/bin*

Biopython

Homepage of the library: http://biopython.org/wiki/Biopython

Version used for testing: *1.54*

Installation with package manager on Debian based systems: *apt-get install python-biopython*

Acknowledgements

Collaborators

We would particularly like to thank SmartArt, who prepared all the graphics and images used within the tool.

Third Party tools

The Genome Analysis Toolkit (GATK) from the Broad Institute is used for the variant calling - an older version (1.6x) is used.

A number of tools from SamTools Picard are also used within the tool, for handling and manipulating SAM and BAM files.

The Pro Data Analysis Module allows access to a number of commonly available, open source bioinformatics tools. These are not however bundled within the Omixon Target application.

The IMGT/HLA database is used for the HLA typing, here are the citations:

- Robinson J, Mistry K, McWilliam H, Lopez R, Parham P, Marsh SGE
  The IMGT/HLA Database
  Nucleic Acids Research (2011) **39 Suppl 1**:D1171-6
- Robinson J, Malik A, Parham P, Bodmer JG, Marsh SGE:
  IMGT/HLA - a sequence database for the human major histocompatibility complex
  Tissue Antigens (2000), **55**:280-287

Tutorials

First Use

The easiest way to get started with Omixon Target is to choose a pre-configured Profile to work with (or more than one). Each Profile already includes a reference and a target.

Fast Start:

- Add Profile
- Import sequencing (fastq) data
- Map the reads and call variants
- Analyse the results

This tutorial only explains the steps required for the Fast Start. There are other tutorials available for using the Expert profile and setting the tool up manually.

Adding a Profile

The first step of the Fast Start involves adding a Profile (the alternative is to create your own 'custom' Profile using Create Profile - see the next tutorial for details).

This function is available from the Data Analysis Dashboard. Adding a new Profile will start a new background task, which (by default) will download and install all the data required to use the Profile. It will also create a new Experiment and Analysis for you, which will be pre-configured with the sequencer chosen, and with the reference and target from within the Profile.

After this Add Profile task has finished (it will take 5 to 10 minutes, depending on the speed of your internet connection), you can move to step 2 of this Fast Start tutorial.

Download problems

If for some reason you cannot download files from within the tool (e.g. due to network security settings) you can download the reference files manually and use the 'Select Local file(s)' option instead of the 'Download file(s)' option in the wizard.

The files for the in-built profiles can be downloaded from here:

BRCA profile:

- http://omixon-download.s3.amazonaws.com/target_ref_hg19_chr13.zip
- http://omixon-download.s3.amazonaws.com/target_ref_hg19_chr17.zip

HLA profile:

- http://omixon-download.s3.amazonaws.com/target_ref_hg19_chr6.zip

CFTR profile:

- http://omixon-download.s3.amazonaws.com/target_ref_hg19_chr7.zip

Data Analysis Example Data

Data Analysis is an optional Module built into Omixon Target.

We offer a few example datasets that go with the in-built profiles. These allow you to easily try out the ‘Map and Align’ and ‘Call Variants’ functions within the tool.

BRCA profile example data – paired illumina data (two fastq files)

http://omixon-download.s3.amazonaws.com/target_brca_example.zip

HLA profile example data – ion torrent data (single fastq file) and paired illumina data (two fastq files)

http://omixon-download.s3.amazonaws.com/target_hla_example.zip

In order to import the fastq you should use the ‘import multiple samples’ function from the Analysis Dashboard (select ‘Go’ on the appropriate Analysis, after Add Profile). For the paired data you need to select the first file (.1.fastq) in the first file selector in the wizard, and the second file (.2.fastq) in the ‘paired data’ file selector. The default pair parameters (orientation, distance) are fine for the example data sets.

Once you have imported the fastq, you will be given the option to run the Map and Align function for the Sample

Import sequencing (fastq) data

Once the Add Profile task from step 1 has finished, you can start to work with the Profile.

The first step is to select a Sample to work with. There is a small data management hierarchy within the tool - the Samples are grouped together with an Analysis, and the Analyses are grouped together within an Experiment. Assuming that you selected to also 'Import Example Data' while adding the profile, there will already be an example Experiment for you to work with.

Select the example Experiment by clicking on it (which will navigate to the Experiment Dashboard), and select the example Analysis in the same way (moves to the Analysis Dashboard). You can click on the Create Sample button in the Analysis Dashboard and create a brand new Sample to work with.

- Select the Sample you wish to work with by clicking on it (moves to the Sample Dashboard). Select the 'Import sequencing data' button, and choose a fastq file (or a pair of files, for paired data) from the file system. Click Finish. This will start a background task. Once this task is finished move to the next step of the tutorial.

Map the reads and call variants

Once you have some fastq data imported, then you can map and align this data plus call variants in a single function.

Either:

Select a Sample with fastq data and from the Sample Dashboard select 'Map and Align' from the buttons on the left or

- Select a Sample (or more than one) from the list of Samples in the Analysis Dashboard and select 'Map and Align' from the buttons on the top
- This will start the Map and Align wizard. You shouldn't need to change any settings in here - you can review the setting and when ready press Finish.

This will start one background task per Sample. Once the background tasks are finished go to the next step.

Results - analysis and visualisation

Once the 'Map and Align' (including variant call) is finished, you will now be able to start analysing the results. You should see two new data items within the Sample - one for the mapped data (in BAM format) and one for the variant calls (in VCF format).

You can visualise the results in the Genome Browser. From the Sample Dashboard select 'Genome browser', or from the Analysis Dashboard select the Sample and then choose 'Browse Sample(s)'. This will open the Genome Browser centered on the target you have defined in the Analysis/Experiment.

You can start to analyse the variants. From the Sample Dashboard select 'Analyse sample variants', which will start a new dashboard-style screen with the same name. In here you can the variants that have been called, in lists that are split by on-target and off-target lists. You can select items in these lists and then 'Jump to Variant' which will move the Genome Browser and show you the variant and short read reads at that position.

From here you can also start the Approval Dashboard (in order to approve or reject variants for further analysis), or you can start the Report Dashboard (to view summary reports for this Sample).

Custom Profile

If none of the built-in Profiles are useful, then you will need to create your own 'custom' Profile within the tool.

The first thing to note is that the Profiles are essentially just convenience wrappers around some Reference sequences and a Target definition. So to create the equivalent of a new Profile all you need to do is to import some reference data (and possibly some known variants for that data), and then create a new Target based on that reference data. Even having a Target is not mandatory, but using one is recommended as there are lots of 'Target-based' functions within the tool.

This tutorial will walk you through the steps needed to do this, with an example.

In this example we will create the profile for one of the 'ready to use' Ion AmpliSeq panels.

Here's the link to the kit:

http://www.invitrogen.com/1/1/12092-ion-ampliseq-cancer-panel.html

Here is a description of a run of the kit, with some test data, if you would like to try it:

http://lifetech-it.hosted.jivesoftware.com/docs/DOC-2180

The amplicons in the kit are documented in a bed file:

http://omixon-download.s3.amazonaws.com/AmpliSeqCancerAmplicons.bed

Option 1: Using the 'Create HG19 Profile' function (recommended)

The fastest way to create a new Profile is to use the one of the two Create Profile functions from the main Data Analysis dashboard. These wizards wrap up the a number of the other steps explained below. There are two options for Create Profile - creating one based on the HG19 reference genome, or creating a 'custom' one with another reference genome.

For the AmpliSeq example, you can do this very simply by using the Create HG19 profile function. This will automatically download and extract all the required reference chromosome data from the Omixon web site.

You should start at the Data Analysis Dashboard.

- Click on 'Create HG19 Profile'
- Fill in the details - Profile Name (e.g. AmpliSeq), Target Name (e.g. AmpliSeq Amplicons), Experiment Name (e.g. 'All my

AmpliSeq data'), Analysis Name (e.g. 'First go'). Press Next.
- Choose one or more sequencers for your Analysis (tip - use CTRL or SHIFT click to select multiple items). Press Next.
- Select Annotation Type - Amplicon. Press Next.
- Select the AmpliSeqCancerAmplicons.bed file from the file system.
- Leave the next step set to 'Download file(s)'.
- Press Finish.

This will start a new background task. This task will take quite a lot of time - it will download most of the HG19 genome and install this reference data into the tool.

The task will parse the bed file and identify which chromosomes it needs from within the bed file. It is important that the bed file uses the standard HG19 notation for chromosome names, i.e. the name has to be 'chr1' and not just '1'. If the chromosome is already configured within the tool then it will not be downloaded again.

Once this task has completed you should see your new Experiment (and Analysis) and you can start importing sequencing data and using your new Target.

Download problems

If for some reason you cannot download files from within the tool (e.g. due to network security settings) you can download the files manually and use the 'Select Local file(s)' option instead of the 'Download file(s)' option in the wizard.

The pre-packaged HG10 zip files can be downloaded from here:

- http://omixon-download.s3.amazonaws.com/target_ref_hg19_chr1.zip
- http://omixon-download.s3.amazonaws.com/target_ref_hg19_chr2.zip
- etc... replace the chr* number with the chromosome you would like to download

Option 2: Using the 'Create Custom Profile' function

If you would like to create a new Profile not based on the HG19 reference genome then this wizard is the easiest way to achieve that.

This wizard is very similar to the Create HG19 Profile wizard above. The main difference is that you can either select an existing reference genome or create a new one during the wizard steps.

You can manually create a new reference genome (see steps below) before starting this wizard, if you wish.

This will start a new background task. Once this task has completed you should see your new Experiment (and Analysis) and you can start importing sequencing data and using your new Target.

Option 3: Manually Creating a Profile

You can also manually recreate the Create Custom Profile steps using the individual functions.

Create a Reference

You should start at the Data Analysis Dashboard.

- Click on 'References' and then click on 'Create New Reference'
- Select the name for your reference and species
- Press 'Finish'

Importing Reference Data

Omixon Target needs individual chromosomes to be added (the whole Human Genome multifasta file will not work correctly within the tool).

You should start at the Data Analysis Dashboard.

- Click on 'References' and then click on the new custom reference you just created in the list.
- Select the 'Import reference data' item in the 'Actions' menu on the left.
- Choose the reference chromosomes - you can select multiple files from the list in one go - e.g. chr1.fa, chr2.fa

This will start a new background task. Once this task has completed you can proceed to the next step.

Importing Known Variants

This is an optional step, but having known variants within the tool can assist with the Variant Calling algorithms.

If you are still in the Reference Genome Dashboard then you are starting in the right place. Otherwise start at the Data Analysis Dashboard, click on 'References' and then click on the new custom reference in the list.

- Select the 'Import known variants' item in the 'Actions' menu on the left.
- Select the matching vcf files for the chromosomes. Note that we need the name of the chromosome (e.g. chr1) to appear in the name of the matching .vcf file - here is an example of one we use within one fo the Profiles: dbsnp-hg19-chr1.vcf.

This will start a new background task. Once this task has completed you can proceed to the next step.

Import Target

You should start at the Data Analysis Dashboard.

- Click on 'Targets' and choose the 'Import Target' button from the menu above the list of Targets.
- Give your new Target a name, e.g. Custom Cancer Gene Panel
- Select the Annotation Type - Amplicon
- Select the bed file for your target regions from the file system.
- Choose the Reference - your new custom reference
- Select the Contigs (chromosomes)
- Press Finish

This will start a new background task. Once this task has completed you can create a new Experiment and start using your new Target.

New Target

The first thing to mention is that you don't **have** to use a Target. However, a number of the features within the tool are built to support the Target concept, and having a well-defined target can lead to better analysis results.

There are two ways to create a new Target within the tool. The first (and easiest) is to import a set of Target annotations using a .gff or .bed file. The second way is to manually import annotations for a reference, and then manually create a Target to use those annotations.

Import Target

You should start at the Data Analysis Dashboard.

- Click on 'Targets' and choose the 'Import Target' button from the menu above the list of Targets.
- Give your new Target a name
- Select the Annotation Type - Amplicon, Exon or Gene
- Select the .bed or .gff file from the file system.
- Choose the Reference
- Select the Contigs
- Press Finish

This will start a new background task.

Manual Target Creation

There are three steps involved with creating a new Target, the first of these is actually performed with the Reference Genome that will be used. For the purposes of this tutorial, it is assumed that the Reference Genome has already been created and configured correctly (see the 'New Reference' tutorial for more on this).

Reference Genome configuration (via the Reference Genome Dashboard):

- Import Annotations (Genes or Exons) 

Target configuration (via Targets lists and Target Dashboard)

- Create Target 
- Configure Target 

Adjusting Memory Usage

When you install Omixon Target, it comes preconfigured with some memory settings. These will vary, depending on which version of Omixon Target you have chosen to install. For the 64 bit versions, the memory is set to 5GB, for the 32 bit versions it is set to 1200MB.

For most analysis, the maximum 5GB memory will be enough. However, it can happen that this is not sufficient memory for the underlying alignment algorithms to run. If you get 'Out of Memory' or 'GC Limit Exceeded' errors, then you will need to either:

- increase the amount of memory available to Omixon Target, or
- run the tool with reduced sensitivity

The first option will be preferable. Running with reduced sensitivity is possible, and saves lots of memory, but doesn't give such good results.

Increase the amount of memory available to Omixon Target

This can be done by editing one of the Omixon Target configuration files. In your Omixon Target installation directory on Windows and Linux you should be able to find a file called ot-gui.vmoptions, and on the Mac it is found in /Applications/OmixonTarget.app/Contents/Info.plist . If you edit this file you should find a single VM ('virtual machine') option which looks like this:

-Xmx5g

Changing this to

-Xmx6g

saving the file and restarting Omixon Target will make 6GB of memory available instead of the standard 5GB.

Run the tool with reduced sensitivity

If you are running with a 32 bit version, and/or you don't have more than 5GB of memory available, the only other option is to reduce the sensitivity of the mapping tool.

This is done by using the 'Advanced Parameters' tab of the 'Map and Align' wizards. In here, you can select a file. You need to create a .txt file called something like advanced.properties (the exact file name doesn't actually matter) and add the following line to the text file:

orm.sampling.default=4

This will cause the mapping tool to sample every 4th position in the reference, rather than every single position. This saves a lot of memory - 60 to 70% of the total. You can also of course try a smaller value, such as:

orm.sampling.default=2

This will save about 30-40% of the memory required.

Dashboards

Most of major functions of Omixon Target are organised into separate Dashboards.

The section provides an overview of the Dashboards and the concepts behind each of these main functions.

Home Dashboard

This is main 'home page' for Omixon Target.

There are only three functions available from this screen:

- Data analysis
- HLA Typing
- Settings

The **Data Analysis** function is for general analysis of sequencing targets, where sequencing data can be mapped and aligned against a reference sequences, and variants can be called, visualised and analysed.

The **HLA Typing** function is only for determining and visualising the HLA types within a set of NGS sequencing data. It is simple, fast and accurate.

The **Settings** function allows you to manage users, licenses and general settings within the tool, plus reset the tool to an 'empty' starting state.

HLA Typing Dashboard

All the HLA Typing functions are available from here.

Firstly, the 'Setup Typing' function should be run. This only needs to be run once.

Once the Setup Typing function has been run then the HLA Typing function can be used. From here, multiple samples can be scheduled for typing simultaneously.

It's possible to view the results from the screen, as well as to delete results from previous runs. Results from multiple analyses can be viewed simultaneously. (Tip - You can use CTRL or SHIFT clicks to select multiple analyses.)

In the header section of the dashboard, you can see the version and date of the IMGT/HLA database that is currently configured and the sequencers which have been set up within the tool.

Settings Dashboard

Reachable from the Home Dashboard, the Settings Dashboard displays an overview of the settings in the tool, and allows access to the user management features, the visual display settings, license settings and the bug report (mail) settings.

There is also the (dangerous!) 'Reset Everything' button. This will cause the whole Application to be reset - i.e. **this will delete all imported and analysed data**, and reset the whole of Omixon Target back to a 'brand new' status, as well as set up the application again

only with the Expert Profile. This function is essentially only recommended for demonstrations, or if you really do want to start again from scratch.

Data Analysis Dashboard

The Data Analysis Dashboard is the starting point for general genomic analysis tasks

The recommended way to start doing data analysis is to use one of the in-built Profiles.

The tool starts with an 'empty' Expert Profile. The first step is either to manually configure this Profile, to add one of the existing pre-configured Profiles to the tool, or to create your own Profile(s). There are three pre-configured profiles available:

- BRCA
- HLA
- CFTR

The pre-configured Profile to be used can be chosen with the 'Add Profile' function. This will start a new background task, which will download and import all the data required to configure the tool for the chosen profile.

In order to create a custom profile, the 'Create HG19 Profile' or the 'Create Custom Profile' functions can be used. With the 'Create HG19 Profile', only a single annotation file (in bed or gff format) is needed in order to configure the profile, the reference data and the known variants can be downloaded and imported automatically. The 'Create Custom Profile' makes it easy to create a profile with a custom (e.g. HG18 or non-human) reference, as it merges three of the other functions within the tool into a single wizard, for convenience ('Import Reference Data', 'Import Known Variants' and 'Import Target').

Once a Profile has been added or created, you can get started by selecting one of the Experiments listed in this Dashboard.

The targets and references for the whole tool can also be configured from here.

Experiment Dashboard

The Experiment is the top level of the small hierarchy organising the genomic data within the tool.

This is a flexible container, and should be used to reflect the nature of your analysis requirements.

An Experiment is a grouping of Analyses. Common configuration elements for a group of Analyses can be specified at the Experiment level, including the references to use, the sequencer, and the targets. This will restrict the Analysis configuration options, and later the options available to the various analysis tasks with the tool.

The Experiment can also be used for analysing and reporting the results of multiple Analyses together.

Analysis Dashboard

The main purpose behind the Analysis is to group together a set of samples.

These Samples will share common configuration within the tool - they will share the same targets, the same references, the same sequencer etc.

The Analysis inherits its configuration from its parent Experiment, and can be used to help split the Experiment into more specific sections. There could be one Analysis for Ion Torrent and one for Illumina data, for example, or one Analysis for one Target (e.g the HLA-A gene) and another for another Target (e.g. the HLA-B gene).

From the Analysis Dashboard, it is possible to browse through and work with individual Samples. It's also possible to select multiple Samples and work with these together, either by launching them in the Genome Browser, or by comparing them using the 'Compare Samples' button, which will launch the Sample Difference screen.

Sample Dashboard

A Sample represents a single set of sequencing data for a single individual.

The Sample has a small life-cycle within the tool:

- Create Sample
- Identify Variants
- Approve Variants

- Approve Sample
- Analyse Approved Sample and Approved Variants

The concept is that the 'raw' variants called may not all be of sufficiently high quality for downstream analysis, and the tool supports a number of Approval functions in order to approve or reject variations and filter them down a set suitable for further analysis. In the case of a HLA Typing analysis, then the downstream analysis step will be to determine the HLA Types of the sample.

Approval Dashboard

Not all variants are interesting. Not all variants can be trusted.

Actual variants that have been discovered during the Variant Call can vary in quality. Some of them may be off-target and therefore not interesting. Some may only have very low coverage and so are not trustworthy. Some may be complex and require manual inspection to figure out what is actually going on in the underlying short read data.

Variants can be 'approved' or 'rejected' in the Approval Dashboard. All variants start in 'pending' status, and there are functions to automatically approve and reject groups of variants by various criteria, and also to manually approve or reject individual variants as well.

Downstream analysis steps will only consider approved variants for their input.

Reference Genome Dashboard

This is where the data for the Reference Genomes is managed.

In here, the actual fasta data for the reference genome (or chromosome) can be imported. Contigs will be automatically created when the fasta file is imported into the tool. We strongly recommend that whole chromosomes are used as the reference sequence, to allow for any off-target reads that might fall on pseudogenes to be properly mapped to the pseudogene and not cause false positives by being incorrectly mapped to a gene.

Known variants from dbSNP can also be imported, in .vcf format. These are important for the quality of the Variant Call, and it's recommended to import them.

It's also possible to import gene and exon annotations, in .vcf format. These can be used for creating new Targets within the tool. If you don't need to create a new Target then you don't need to worry about this function.

Target Dashboard

Targets can be configured within this dashboard.

Targets are optional within the tool, but are highly recommended, as there are a number of analysis features that give better results for well-defined targets.

A typical Target could be 'all the exons within the BRCA1 and BRCA2 genes'.

At the moment, the target configuration is restricted to gene, exon and amplicon annotations. There are two ways to import annotations - either via the 'Import Target' method, where annotations can be imported and a Target created in a single step, or the annotations to be used can be imported via the 'Import Annotations' function in the Reference Genome Dashboard - this second method requires a Target to be created or altered manually in order to make use of these annotations.

Other Screens

In addition to the dashboards there are some other function-rich screens that are mostly used for visualising and manipulating the results of the analysis tasks.

Analyse Sample Variants

This screen is the starting point for the analysis of the variants found in a Sample.

If there are known ('expected') variants already found for this sample, these can be compared with the actual variants found during the analysis.

The variants found are listed according to whether they are on or off target, and it's possible to select individual variants and jump into the Genome Browser to inspect them.

The Quality Control results are also summarised on this screen.

It's possible to start the Approval Dashboard from here, in order to approve the variants ready for further analysis (and reject any low quality or off target variants that should be ignored).

It's also possible to navigate to the Report Dashboard to see graphs, charts and other summary information about the Sample data.

Sample Difference

This is the starting point for multiple sample comparison, for example with trio data (mother, father, child).

This screen is started by selecting multiple Samples within the Analysis Dashboard, and then selecting 'Compare Samples'.

Variants within the comparison table can be selected, and then you can 'jump' to that position within the Genome Browser, which will also allow you to browse all the samples being viewed in the Samples Difference screen together.

Genome Browser

The Genome Browser allows visual inspection of genomics data.

Multiple Samples can be browsed together.

The data items that can be visualised include:

- Reference Sequence
- Gene Annotations
- Exon Annotations
- Amplicon Annotations

And for each Sample:

- Actual Variant Annotations
- Expected Variant Annotations
- Short Reads, including Coverage

By default, the display is 'masked' so that only differences between the short reads and the reference sequence are displayed. This mask is done in a stranded fashion, so that forward strand reads (pink) and reverse strand reads (yellow) can be easily distinguished within the display. The 'Display Strand' function turns this on and off.

The short reads track can be 'collapsed' which gives a summary view of the short reads (and does not allow each read to be inspected in detail).

Individual items (annotations, short reads) within the Genome Browser can be selected, and it's possible to browse elsewhere and then jump back to the selected item, or copy the details of the selected item to the clipboard.

Expanding the display

At the top left of the tool is a small control that allows an expanded view. This actually works for all screens in the tool, but is most useful for the Genome Browser. Clicking the same icon again will go back to the usual display.

Rotating the Genome Browser

The vertical Genome Browser view is more useful for comparing multiple samples. The view can be rotated to use a horizontal display, which is a bit more traditional for Genome Browsers, and is more useful when browsing a single sample.

Zooming

By default the Genome Browser starts in 'Drag' mode, where the mouse wheel (or '+' and '-' keys) can be used to zoom in and out, and the display can be moved by clicking and dragging with the mouse. 'Drill' mode can also be used, where the region highlighted by a mouse click and drag will be 'drilled into' once the mouse button is released.

Jumping Around

You can jump to a position in the reference using 'Jump To Position', or jump to one of the annotations within the Target using 'Jump To Target'. If you have previously selected a feature within the display (such as an annotation or short read), then you can jump back to that feature at any time by using the small arrow button at the bottom of the display.

Filtering the Short Reads

You can display more short reads by using the 'collapsed' pile-up view.

You can also 'page' around the short reads using the small arrows at the top of the short read track. You can set the page size using the up and down arrows, and the current page by using the left and right arrows. This is useful when you have deep sequencing and want to scroll through a few hundred short reads at a time.

Filtering the Variants

The variants in the 'Act' (Actual) variants track can also be filtered. This makes use of the 'Approval' features within the tool - individual

variants can be 'approved' or 'rejected' (or still be 'pending', awaiting either approval or rejection). This filter will display 'All' variants (three lines, default), then 'Pending' (circle), 'Approved' (tick) or 'Rejected' (cross).

Track Setup

You can configure which tracks you would like to see in the display.

Exporting Genome Browser Data

You can 'Copy to Clipboard' the details of individual, selected items in the display.

You can also 'Capture a Screenshot' of the currently visibly tracks within the Genome Browser. This will create a .png file at your chosen location.

Analysis Results

Summary statistics for all the Samples in the Analysis

This is a very simple summary screen that is available from the Analysis Results button in the Analysis Dashboard. It displays overview statistics of the progress for all the Samples in the Analysis.

References

Centralised reference configuration for the whole tool

Omixon Target has a centralised configuration for the References (reference sequences) used in the Data Analysis portion of the tool.

These References can be set up once, and then used in multiple Experiments and Analyses.

This screen lists the available Reference Genomes, and allows Reference Genomes to be created or deleted.

Clicking on a Reference Genome will take you to the Reference Genome Dashboard, where reference data (in fasta format) can be imported, along with known variants (from dbsnp).

Targets

Centralised Target configuration for the whole tool.

Omixon Target has a centralised configuration for the Targets (essentially lists of annotations) used in the Data Analysis portion of the tool.

These Targets can be set up once, and then used in multiple Experiments and Analyses. Targets are actually optional, but recommended.

This screen lists all the Targets defined, and allows Targets to be created or deleted. It is possible to create a Target by importing annotations in BED or GFF format using the 'Import Target' function.

Clicking on a Target will take you to the Target Dashboard, where the Target can be configured.

Experiment References

Configure References for an Experiment

Omixon Target has a centralised configuration for the References (reference sequences) used in the Data Analysis portion of the tool.

These References can be set up once, and then used in multiple Experiments and Analyses.

This screen lists the Reference Genomes linked to this Experiment, and allows Reference Genomes to be created or deleted.

Clicking on a Reference Genome will take you to the Reference Genome Dashboard, where reference data (in fasta format) can be imported, along with known variants (from dbsnp).

Experiment Targets

Configure Targets for an Experiment

Omixon Target has a centralised configuration for the Targets (essentially lists of annotations) used in the Data Analysis portion of the tool.

These Targets can be set up once, and then used in multiple Experiments and Analyses.

This screen lists the Targets attached to this Experiment, and allows Targets to be created or deleted. It is possible to create a Target by importing annotations in BED or GFF format using the 'Import Target' function.

Clicking on a Target will take you to the Target Dashboard, where the Target can be configured.

Analysis References

Configure References for an Analysis

Omixon Target has a centralised configuration for the References (reference sequences) used in the Data Analysis portion of the tool.

These References can be set up once, and then used in multiple Experiments and Analyses.

This screen lists the Reference Genomes linked to this Analysis, and allows Reference Genomes to be created or deleted.

Clicking on a Reference Genome will take you to the Reference Genome Dashboard, where reference data (in fasta format) can be imported, along with known variants (from dbsnp).

Analysis Targets

Configure Targets for an Analysis

Omixon Target has a centralised configuration for the Targets (essentially lists of annotations) used in the Data Analysis portion of the tool.

These Targets can be set up once, and then used in multiple Experiments and Analyses.

This screen lists the Targets attached to this Analysis, and allows Targets to be created or deleted. It is possible to create a Target by importing annotations in BED or GFF format using the 'Import Target' function.

Clicking on a Target will take you to the Target Dashboard, where the Target can be configured.

HLA Typing Analysis Result

A high-level overview of the HLA Haplotying results

A high-level overview of the HLA Typing results for a single Analysis (i.e. for a single submission of the HLA Typing tool) or for multiple Analyses. This can include the summary results for many samples.

Each Sample can be selected and the detailed results for a sample or an allele, or a short read alignment visualisation can be viewed for that Sample, by pressing the **Sample Details**, the **Allele Details** or the **Browse Sample** buttons, respectively.

A minimum coverage threshold can be set for the summary result table. This limit can be set as an absolute coverage value (i.e. the mean number of reads for the allele) or a relative coverage threshold (minimum coverage relative to the top allele). The default value for minimum coverage is 10 for the absolute and 95% for the relative coverage filter function. Alleles below the coverage limit are shown as empty cells. Note, that allele candidates with a lower coverage than the selected limit are NOT shown on the "HLA Typing sample result" dashboard, which contains the detailed results.

For each allele, one or more allele candidates can be assigned. There are four different ways to assign allele candidates:

1. Candidates can be assigned manually by clicking on the checkmark before the allele candidate's name.
2. All 'winner' candidates (i.e. the best allele candidates based on coverage statistics) can be assigned using the **Assign Winners** button.
3. All unambiguous results can be assigned using the **Assign Unambiguous** button.
4. All allele candidates can be assigned using the **Assign All** button.

All assignments can be deleted by the **Unassign All** button. Note, that the 'Assign ...' buttons on this screen affect all the displayed samples, if you prefer to do the assignments separately for each sample/allele, you can find similar functions on the HLA Typing sample and allele result screens.

The results can be exported in TXT (tab delimited text), CSV (comma separated text) or XLS (Excel) format.

Tip - If you would like to export only the assigned allele candidates click the **Assigned Only** button, then export the results.

At this level, the summary HLA Typing results can be displayed in 4-, 6- or 8-digit formats. In the more detailed results always the 6- or 8-digit format can be seen.

You can filter the displayed loci by using the **Setup Loci** function.

HLA Typing Sample Result

A more detailed view of the HLA Typing results for a single Sample.

For each allele at each locus we supply:

- Detection (%) - A percentage value, based on a comparison of the density of coverage. The allele uses the lowest detection value among its exons. High values are better.
- Average coverage - The average number of short reads covering the whole exon or allele (exons with zero coverage are not counted).
- Exons covered - Number of exons with non-zero coverage / number of available exon sequences for the allele candidate.

By clicking on the Browse button next to the name of a specific allele, the **HLA genome browser** can be opened. This HLA visualisation tool shows the short reads in the Sample aligned to the allele candidates. The best allele candidate is marked as "Winner".

By clicking on the Analyse button next to the allele name, the **HLA Typing allele result** screen is opened. This page contains detailed, exon level statistics for the allele candidates.

A minimum coverage threshold can be set for the summary result table. This limit can be set as an absolute coverage value (i.e. the mean number of reads for the allele) or a relative coverage threshold (minimum coverage relative to the top allele). The default value for minimum coverage is 10 for the absolute and 95% for the relative coverage filter function. Alleles below the coverage limit are shown as empty cells. Note, that allele candidates with a lower coverage than the selected limit are NOT shown on the "HLA Typing sample result" dashboard, which contains the detailed results.

For each allele, one or more allele candidates can be assigned. There are four different ways to assign allele candidates:

1. Candidates can be assigned manually by clicking on the checkmark before the allele candidate's name.
2. All 'winner' candidates (i.e. the best allele candidates based on coverage statistics) can be assigned using the **Assign Winners** button.
3. All unambiguous results can be assigned using the **Assign Unambiguous** button.
4. All allele candidates can be assigned using the **Assign All** button.

All assignments can be deleted by the **Unassign All** button. Note, that the 'Assign ...' buttons on this screen affect all the displayed samples, if you prefer to do the assignments separately for each sample/allele, you can find similar functions on the HLA Typing sample and allele result screens.

The results can be exported in TXT (tab delimited text), CSV (comma separated text) or XLS (Excel) format.

Tip - If you would like to export only the assigned allele candidates click the **Assigned Only** button, then export the results.

You can filter the displayed loci by using the **Setup Loci** function.

HLA Typing Allele Result

A more detailed view of the HLA Typing results for a single Allele.

For each allele candidate we supply:

- Detection (%) - A percentage value, based on a comparison of the density of coverage. The allele uses the lowest detection value among its exons. High values are better.
- Average coverage - The average number of short reads covering the whole exon or allele (exons with zero coverage are not counted).
- Exons covered - Number of exons with non-zero coverage / number of available exon sequences for the allele candidate.
- "Detection" and "Average coverage" values for each covered exons.

A minimum coverage threshold can be set for the allele result table. This limit can be set as an absolute coverage value (i.e. the mean number of reads for the allele) or a relative coverage threshold (minimum coverage relative to the top allele). The default value for minimum coverage is 10 for the absolute and 95% for the relative coverage filter function. Alleles below the coverage limit are shown as empty cells. Note, that allele candidates with a lower coverage than the selected limit are NOT shown on the "HLA Typing allele result" dashboard, which contains the detailed results.

For each allele, one or more allele candidates can be assigned. There are four different ways to assign allele candidates:

1. Candidates can be assigned manually by clicking on the checkmark before the allele candidate's name.
2. All 'Winner' candidates (i.e. the best allele candidates based on coverage statistics) can be assigned using the **Assign Winners** button.
3. All unambiguous results can be assigned using the **Assign Unambiguous** button.
4. All allele candidates can be assigned using the **Assign All** button.

All assignments can be deleted by the **Unassign All** button.

The results can be exported in TXT (tab delimited text), CSV (comma separated text) or XLS (Excel) format.

Tip - If you would like to export only the assigned allele candidates click the **Assigned Only** button, then export the results.

By clicking on the **Browse results** button, the **HLA genome browser** can be opened. This HLA visualisation tool shows the short reads in the Sample aligned to the allele candidates. The best allele candidate is marked as 'Winner'.

HLA Genome Browser

The HLA Genome Browser allows visual inspection of genomics data.

Multiple allele candidates can be browsed together.

The data items that can be visualised for each allele candidate include:

- Reference Sequence
- Exon Annotations
- Short Reads
- Coverage

By default, short read alignments are displayed in a stranded fashion, so that forward strand reads (pink) and reverse strand reads (yellow) can be easily distinguished within the display. The 'Display Strand' function turns this on and off.

The short reads track can be 'collapsed' which gives a summary view of the short reads (and does not allow each read to be inspected in detail).

Individual items (exons, short reads) within the HLA Genome Browser can be selected, and it's possible to browse elsewhere within the same locus and then jump back to the selected item, or copy the details of the selected item to the clipboard.

Expanding the display

At the top left of the tool is a small control that allows an expanded view. This actually works for all screens in the tool, but is most useful for the HLA Genome Browser. Clicking the same icon again will go back to the usual display.

Rotating the Genome Browser

The vertical HLA Genome Browser view is more useful for comparing multiple allele candidates. The view can be rotated to use a horizontal display, which is a bit more traditional for Genome Browsers, and is more useful when browsing a single allele candidate.

Zooming

By default the Genome Browser starts in 'Drag' mode, where the mouse wheel (or '+' and '-' keys) can be used to zoom in and out, and the display can be moved by clicking and dragging with the mouse. 'Drill' mode can also be used, where the region highlighted by a mouse click and drag will be 'drilled into' once the mouse button is released.

Jumping Around

You can jump to a position in the reference using 'Jump To Position'. If you have previously selected a feature within the display (such as an annotation or short read), then you can jump back to that feature at any time by using the small arrow button at the bottom of the display.

Filtering the Short Reads

You can display more short reads by using the 'collapsed' pile-up view.

You can also 'page' around the short reads using the small arrows at the top of the short read track. You can set the page size using the up and down arrows, and the current page by using the left and right arrows. This is useful when you have deep sequencing and want to scroll through a few hundred short reads at a time.

Filtering the Allele Candidates

The list of allele candidates shown in the HLA Genome browser can be filtered by coverage. This limit can be set as an absolute coverage value (i.e. the mean number of reads for the allele) or a relative coverage threshold (minimum coverage relative to the top allele). The default value for minimum coverage is 10 for the absolute and 95% for the relative coverage filter function. Current coverage filter settings can be seen in the top right corner, above the alignments.

By clicking the 'Assigned Only' button, unassigned allele candidates can be hidden.

Analysing result

By clicking on the 'Analyse Result' button, the HLA Typing allele result screen is opened. This page contains detailed, exon level statistics for the allele candidates. Note, that only samples visible in the current browser session will be displayed in the allele result table.

Track Setup

You can configure which tracks you would like to see in the display.

Display Setup

You can display/hide read pair information, indels, SNPs and soft-clips.

Exporting HLA Genome Browser Data

You can 'Copy to Clipboard' the details of individual, selected items in the display.

You can also 'Capture a Screenshot' of the currently visibly tracks within the HLA Genome Browser. This will create a .png file at your chosen location.

User Management

User Management allows Users to be added, edited and removed from the system.

There must be at least one 'Super User' at all times. The first registered user automatically becomes the Super User, this user cannot be deleted.

Configure Application

This wizard allows you to set up a mail server via SMTP, for sending bug and error reports directly to Omixon.

You can adjust the colour scheme of the Genome Browser here.

The "zooming resolution" of the Genome Browser can also be set on this screen by changing the 'Overview factor'. This is a the maximum number of nucleotides displayed per pixel. The default setting is 20.

Settings selected in the Configure Application wizard are saved and will still be in effect after Omixon Target is restarted.

Wizards

"Wizards" are dialogs that pop up on top of the main display area in Omixon Target, and allow the user to choose parameters and options for various processes.

In all cases, as many options as possible are already prefilled and preselected within the wizards. If only one Target, one Sequencer, one Reference Genome etc are defined, then there will be no need to select these again within the wizard.

In many cases, there will not be any need to change any options within the wizards at all - they can be started and the Finish button hit with no other interaction required.

Upload license

You can upload a license with this wizard. Just select the file you've got from Omixon and finish the wizard.

The license can be an upgrade from an evaluation license to a full license (either permanent license or annually renewable), or could just be a top-up license for the credit based services (such as HLA Typing).

HLA Typing Setup

This only needs to be run once (unless you would like to use another sequencer).

- Choose the sequencer(s) you would like to use, and hit finish. (Tip - you can use CTRL and SHIFT clicks to select multiple sequencers.)

A configuration file will be downloaded from the internet and installed into Omixon Target. This should only take a few seconds.

Once this has finished you can start to use the HLA Typing function.

Download problems

If for some reason you cannot download the HLA configuration file from within the tool (e.g. due to network security settings) you can download the files manually and use the 'Select Local file(s)' option instead of the 'Download file(s)' option in the wizard.

The in-built HLA configuration file can be downloaded from here:

- http://omixon-download.s3.amazonaws.com/target_haplotype_db.zip

HLA Typing

This function is available from the HLA Typing Dashboard. Before running HLA Typing, the Setup Typing wizard will need to be run.

You need to fill in a name for this HLA typing Analysis, plus select fastq files for the input (you can select multiple files, and multiple paired files). You also need to select species and which sequencer was used to created the fastq data.

This function will allow you to process multiple fastq files - or multiple pairs of matching fastq files - at the same time. The second file is optional, and is for paired reads - i.e. if your analysis produced paired end or mate pair data. If using paired reads, the two input files are assumed to have the exact matching reads, in the exact same order (there is currently no in-built check for this).

You can also use SFF files as input or you might combine FASTQ and SFF files as well. Note that if any of the input files is in format SFF then paired import mode is not supported.

A background task will be started, and once it's finished a new item will appear in the main list within the HLA Typing Dashboard where you will be able to see the results of the HLA Typing.

Note that the typing process has different categories and which could be executed depending on your license credits. The license category can be selected on the first page of the wizard.

Paired Read Options

- Minimum Distance - the minimum expected distance between the two reads
- Maximum Distance - the maximum expected distance between the two reads
- Orientation - the orientation of the reads with respect to each other. The first file is the first in the orientation as well. Options are FR (forward-reverse, i.e. first file contains forward reads, second file contains reverse reads), RF (reverse-forward) and FF (forward-forward).

Advanced Options

These options dicatate where the data has come from and effect how the underlying algorithms deal with the data. Some data sources (e.g. whole genome and whole exome) are more 'noisy', and the algorithms can help to filter out this extra noise.

- Choose what the source is of the HLA Typing data.
- Maximum Reads Processed. **This is very important.** This dictates how many reads (or how many pairs) will be processed from each input file in order to do the HLA Typing. You can now choose from a pre-configured set of typical sequencing runs, including Whole Genome and Whole Exome. You can also choose Custom and manually set how many reads to process. If you are not sure, then use the 'All' option to process all the reads. For very targeted data sets including only the HLA loci, this value doesn't need to be too high (about 10'000). For larger kits such as the RainDance HLAseq kit (the whole MHC) this will need to be significantly higher (at least 2'000'000). For whole exome or whole genome data sets it is recommended to process all the reads (or pairs), because of the lower coverage that is usually found in these kind of data sets.

Add Profile

The tool starts up in '**Expert** mode', with only the Expert Profile configured and no reference data imported or targets set up.

A 'Profile' is simply a wrapper around a Target, plus a set of reference data and known variants.

It is possible to add a pre-configured Profile to the tool using the 'Add Profile' function.

The available pre-configured profiles are:

- **HLA** (chr6)
- **BRCA** (chr13 and chr17)
- **CFTR** (chr7)

The three pre-configured profiles can be set up automatically by using this function, which will also download and import all the data required for each Profile.

You can choose not to add a pre-configured Profile and create your own using the 'Create Profile' option. It's also possible to skip Profile set-up entirely, and simply configure the tool manually.

Download problems

If for some reason you cannot download the profile configuration file from within the tool (e.g. due to network security settings) you can download the necessary files manually and use the 'Select Local file(s)' option instead of the 'Download file(s)' option in the wizard.

For each profile, the setup files are packaged by chromosome and can be downloaded from the following links:

BRCA profile:

- http://omixon-download.s3.amazonaws.com/target_ref_hg19_chr13.zip

- http://omixon-download.s3.amazonaws.com/target_ref_hg19_chr17.zip

HLA profile:

- http://omixon-download.s3.amazonaws.com/target_ref_hg19_chr6.zip

CFTR profile:

- http://omixon-download.s3.amazonaws.com/target_ref_hg19_chr7.zip

Create HG19 Profile

The tool starts up in 'Expert mode', with only the Expert Profile configured and no reference data imported or targets set up.

A 'Profile' is simply a wrapper around a Target, plus a set of reference data and known variants.

A Target is simply a list of reference annotations, for either genes, exons or amplicons.

It is possible to add a pre-configured Profile to the tool using the 'Add Profile' function, or you can use this 'Create HG19 Profile' option instead. It's also possible to skip Profile set-up entirely, and simply configure the tool manually.

- Fill in the details - Profile Name, Target Name, Experiment Name, Analysis Name. A new Profile, Target, Experiment and Analysis will be created. Press Next.
- Choose one or more sequencers for your Analysis (tip - use CTRL or SHIFT click to select multiple items). Press Next.
- Select Annotation Type - Exon/Gene/Amplicon. Press Next.
- Select a .bed or .gff file from the file system, containing the annotations that describe your Target. Press Finish.

Create Profile

The tool starts up in 'Expert mode', with only the Expert Profile configured and no reference data imported or targets set up.

A 'Profile' is simply a wrapper around a Target, plus a set of reference data and known variants.

A Target is simply a list of reference annotations, for either genes, exons or amplicons.

It is possible to add a pre-configured Profile to the tool using the 'Add Profile' function, or you can use this 'Create Profile' option instead. It's also possible to skip Profile set-up entirely, and simply configure the tool manually.

- Fill in the details - Profile Name, Target Name, Experiment Name, Analysis Name. A new Profile, Target, Experiment and Analysis will be created. Press Next.
- Choose one or more sequencers for your Analysis (tip - use CTRL or SHIFT click to select multiple items). Press Next.
- Select a reference. Press Next.
- Choose the reference files. Press Next.
- Optional Step - select the matching Known Variants vcf files for the reference files. Note that we need the name of the contig (e.g. chr17) to appear in the name of the matching .vcf file - here is an example of one we use within one of the Profiles: dbsnp-hg19-chr17.vcf. Press Next.
- Select Annotation Type - Exon/Gene/Amplicon. Press Next.
- Select a .bed or .gff file from the file system, containing the annotations that describe your Target. Press Finish.

Import Multiple Samples

This function is available from the Analysis Dashboard.

If you have short read data from sequencing run you can import it with this function, and then run the 'Map and Align' function, which will align the data to a Reference Genome and run a variant call in order to identify the variants.

It is essentially identical to the 'Import Sequencing Data' function within the Sample Dashboard - the goal it to import single or multiplexed fastq/sff data into the tool. The difference with this version is that it will do two things:

- Import fastq/sff data (or paired fastq data)
- Automatically create a new Sample for each imported fastq file (or pair of fastq files) in case of single sample data
- Demultiplex each fastq/sff file and create a new Sample for each sample found in the files identified by barcodes

This function will allow you to import multiple fastq/sff files - or multiple pairs of matching fastq files - at the same time. The second file is optional, and is for paired reads - i.e. if your analysis produced paired end or mate pair data. If using paired reads, the two input files are assumed to have the exact matching reads, in the exact same order (there is currently no in-built check for this). The paired read import is not supported for sff files. Note that the sff import function is only available in OmixonTarget Pro edition.

Steps in this wizard:

- Select Input File
- Select Paired Read Options
- Demultiplexing Options (Omixon Target Pro edition only)

After you have imported sequencing data, you can align it and call variant in the data using the Map and Align function.

Paired Read Options

- Minimum Distance - the minimum expected distance between the two reads
- Maximum Distance - the maximum expected distance between the two reads
- Orientation - the orientation of the reads with respect to each other. The first file is the first in the orientation as well. Options are FR (forward-reverse, i.e. first file contains forward reads, second file contains reverse reads), RF (reverse-forward) and FF (forward-forward).

Demultiplexing Options (Omixon Target Pro edition only)

- Sequencer - the sequencer used to produce the multiplexed data
- Barcode file - text file containing barcode names and related sequences to identify the samples in the multiplexed files

Note that Illumina demultiplexing is supported only for paired fastq files and Roche 454 demultiplexing is supported only for not paired sff files. Ion Torrent data demultiplexing is not supported currently.

Create Sample

This wizard allows you to create an empty Sample. You need to give it a name (which should be unique within the Analysis) and select a species. Once a Sample is created, data can be imported for analysis.

The other way to create Samples is to use the 'Import multiple samples' function from the Analysis Dashboard, which will both import fastq and automatically create Samples at the same time.

Map and Align Samples

If you have imported sequencing data, then you can map and align this data to a Reference Genome using this function.

This version of the Map and Align runs with multiple samples - i.e. all the samples selected in the list visible in the Analysis Dashboard. If any of the samples selected already have mapped data and/or variant calls they will be (silently) ignored.

This function has three analyses built in to it:

- Map and Align using Omixon mappers
- Variant Call using GATK (optional)
- Quality Control (coming soon!)

The Map and Align step will always be performed. The Variant Call and Quality Control steps are optional. The Variant Call can be performed later using the 'Call Variants' function, or the mapped data can be copied elsewhere and other variant caller used instead. The Quality Control cannot be run separately, only as part of the Map and Align process.

The steps in this wizard:

- Select Reference Data
- Select Species
- Select Sequencer
- Advanced Options
- Variant Call Options
- (Quality Control Options - coming soon!)

Like all the wizards within the tool, most of the options within this wizard will already be pre-filled and pre-selected, based on your chosen Analysis and Experiment configuration.

Advanced Options

- Parameters file: You can choose to import a properties file with some advanced parameters. This file is the same one used for the Omixon Variant Toolkit (the command line version of our alignment algorithms), you can find the latest readme file for this linked via the Omixon web site (the readme file link is in the 'Useful Links' section at the bottom of the page): https://www.omixon.com/omixon/abouttoolkit.htm
- Maximum Coverage: You can choose what the maximum depth of coverage should be for your results. For very deep coverage data, it's usually enough to keep 1000 deep short reads, however if you want to you can keep more or less than this. This will cause your actual mapping results to be discarded, only a number matching your maximum coverage will be kept.

Variant Call Options

You can choose whether or not to run a GATK variant call after the map and align function has run. You can always run the variant call later, if desired.

- The dbSNP (known variants) file is an optional parameter. You can either use dbSNP data that has been imported into the tool, or select a dbSNP file from the file system.

Import Sequencing Data

The first goal of Sample analysis is to identify some variants. There are three ways to do this, and this function allows one of these three ways.

If you have short read data from sequencing run you can import it with this function, and then run the 'Map and Align' function, which will align the data to a Reference Genome and run a variant call in order to identify the variants.

You will usually either run this wizard from within an existing Sample, or use the Import Multiple Samples wizard from the Analysis Dashboard, which both imports data and creates (multiple) samples at the same time.

This function will allow you to import two FASTQ files at the same time (or one SFF). The second one is optional, and is for paired reads if your analysis produced paired end or mate pair data. If using paired reads, the two input files are assumed to have the exact matching reads, in the exact same order (there is currently no in-built check for this). The SFF import doesn't support paired mode and is only available in Omixon Target Pro.

Steps in this wizard:

- Select Input File
- Select Paired Read Options

After you have imported sequencing data, you can align it and call variant in the data using the Map and Align function.

Paired Read Options

- Minimum Distance - the minimum expected distance between the two reads
- Maximum Distance - the maximum expected distance between the two reads
- Orientation - the orientation of the reads with respect to each other. The first file is the first in the orientation as well. Options are FR (forward-reverse, i.e. first file contains forward reads, second file contains reverse reads), RF (reverse-forward) and FF (forward-forward).

Import Mapped Data

The first goal of Sample analysis is to identify some variants. There are three ways to do this, and this function allows one of these three ways.

If you have short read data that you have already mapped from another source you can skip the 'Import Sequencing Data' and 'Map and Align' functions altogether, and just import your mapped data into the tool (in either .sam or .bam format), and then use the Call Variants function to identify the variants.

This wizard also allows you to import Sanger data, however the Sanger visualisation is currently under development.

Steps in this wizard:

- Select Mapping Type
- Select Input File
- Select Reference Contigs

After you have imported mapped data, you can browse it via the Genome Browser, or use the Call Variants function to identify variants within the mapped data.

Import Variant Calls

The first goal of Sample analysis is to identify some variants. There are three ways to do this, and this function allows one of these three ways.

If you have already mapped your data and called variants on it, or if you have variants already identified from another source (such as Sanger sequencing), then you can skip the 'Map and Align' and 'Call Variants' functions altogether and simply import your variants into the tool.

This function also allows you to import 'expected' variants. If you are working with simulated data, or if you already have known variants from another source then you can import these and tag them as 'expected' variants. When you have then identified variants (via an import of 'actual' variants, or via a Map and Align or Call Variants analysis), then you can automatically compare the expected variants with the actual variants using the Analyse Sample Variants screen.

Steps in this wizard:

- Select Annotation Type
- Select Input Files
- Import Options
- Select Contig

After you have imported variant calls, you can visualise them in the Genome Browser, analyse them within the Analyse Sample Variants screen, and compare them with the variant calls of other Samples using the Compare Samples function.

Import Sanger

This function allows you to import multiple Sanger sequences, for visualisation alongside your short reads in the Genome Browser.

This function is currently under development, however you should already be able to see multiple traces for the Sanger data.

Map and Align

If you have imported sequencing data, then you can map and align this data to a Reference Genome using this function. This version of the Map and Align runs with a single sample. This function has three analyses built in to it:

- Map and Align using Omixon mappers
- Variant Call using GATK (optional) or Samtools (only in Target Pro edition)
- Quality Control (coming soon!)

The Map and Align step will always be performed. The Variant Call and Quality Control steps are optional. The Variant Call can be performed later using the 'Call Variants' function, or the mapped data can be copied elsewhere and other variant caller used instead. The Quality Control cannot be run separately, only as part of the Map and Align process.

The steps in this wizard:

- Select Genomic Data
- Select Species
- Select Sequencer
- Advanced Options
- Variant Call Options
- (Quality Control Options - coming soon!)

Like all the wizards within the tool, most of the options within this wizard will already be pre-filled and pre-selected, based on your chosen Analysis and Experiment configuration.

Advanced Options

- Parameters file: You can choose to import a properties file with some advanced parameters. This file is the same one used for the Omixon Variant Toolkit (the command line version of our alignment algorithms), you can find the latest readme file for this linked via the Omixon web site (the readme file link is in the 'Useful Links' section at the bottom of the page): https://www.omixon.com/omixon/abouttoolkit.htm
- Maximum Coverage: You can choose what the maximum depth of coverage should be for your results. For very deep coverage data, it's usually enough to keep 1000 deep short reads, however if you want to you can keep more or less than this. This will cause your actual mapping results to be discarded, only a number matching your maximum coverage will be kept.

Variant Call Options

You can choose whether or not to run a GATK variant call after the map and align function has run. You can always run the variant call later, if desired.

- The dbSNP (known variants) file is an optional parameter. You can either use dbSNP data that has been imported into the tool, or select a dbSNP file from the file system.

In the OmixonTarget Pro edition, you can also choose to run a Samtools mpileup variant call instead of a GATK variant call. This is recommended for 454 and IonTorrent data.

Call Variants

If the Map and Align function is used to map short reads, then there is an option to run a Variant Call within that wizard. If that option is not chosen, or if already mapped reads are imported into the tool, then a separate Variant Call can be started.

The Variant Call invokes a full GATK pipeline, following the recommended best practises of the Broad Institute.

Like all wizards in the tool, as many options as possible are already preselected within the wizard.

These are the wizard steps:

- Select Genomic Data
- Select Species
- Select Sequencer
- Advanced Options
- Variant Call Options

Advanced Options

- Parameters file: This is currently not used for the variant caller.

Variant Call Options

- The dbSNP (known variants) file is an optional parameter. You can either use dbSNP data that has been imported into the tool, or select a dbSNP file from the file system.

In Omixon Target Pro edition, you can also choose to run a Samtools mpileup based variant call instead of a GATK pipeline. This is recommended for 454 and IonTorrent data.

Auto Approve

Even with targeted sequencing, there can be a large number of variants to deal with.

Auto approve allows the 'bulk' approval of variants that meet one or more of a number of criteria, including:

- The quality of the call is above a minimum quality
- The coverage at the location is greater than a minimum coverage
- The call is on Target

Auto Reject

Even with targeted sequencing, there can be a large number of variants to deal with.

Auto reject allows the 'bulk' rejection of variants that **fail** to meet one or more of a number of criteria, including:

- The quality of the call is above a minimum quality
- The coverage at the location is greater than a minimum coverage
- The call is on Target

Auto Reset

If you have made a mistake with one of the Auto Accept or Auto Reject functions, you can reset all your variants back to 'pending' status by using the Auto Reset function. This is an all-or-nothing operation.

Design Analysis

This wizard usually only has to be run once, when a new Analysis is created.

This wizard inherits all its settings from the parent Experiment, and can be used to restrict the options from the Experiment for each Analysis.

The steps in the wizard:

- General Properties
- Choose Sequencer
- Choose Target
- Choose Species
- Choose Genome
- Choose Contig

Once all these properties are set for an Analysis, there will be automatically used by all the wizards and processes for the Samples that belong to the Analysis.

Design Experiment

This wizard usually only has to be run once, when a new Experiment is created.

This wizard inherits all its settings from the Profiles have been added to the tool, and the lists of configured Reference Genomes and Targets.

The steps in the wizard:

- General Properties
- Choose Sequencer
- Choose Target
- Choose Species
- Choose Genome
- Choose Contig

Once all these properties are set for an Experiment, there will be automatically used by Analyses that belong to the Experiment.

Create Reference

This wizard is used to create a new Reference Genome within Omixon Target. Once created, this Reference will become available for all Experiments and Analyses to use.

The options in this wizard:

- Choose Reference Name
- Select Species

It is not sufficient to simply create a Reference Genome - before it can be used with in the tool it should also have reference data, known variants and gene/exon annotations imported.

Import Reference Data

This function will import the reference data for the Reference Genome.

It is recommended to import and use whole chromosomes. Even targeted sequencing can place some reads off target, and it's very important that off target reads are also mapped off target correctly, particularly if they happen to map to a pseudogene of the gene(s) being targeted.

The input file must be in fasta format. We recommend using multiple fasta files within Omixon Target, rather than Multifasta references.

Contigs will be automatically created, based on the names found within the input fasta file.

Import Known Variants

This function will import dbSNP known variant annotations for the Reference Genome.

This will be used later with the GATK variant call.

If the Reference Genome has been configured to use a whole chromosome, then we recommend importing all matching dbSNP records for that chromosome.

The annotations must be in .vcf format.

Import Annotations

This function will import Gene or Exon annotations for the Reference Genome.

This will be used later for Target configuration.

Currently, the annotations must be in .bed or .gff format. If you need another format please let us know.

The steps in this wizard are:

- Select Annotation Type
- Select Input Files
- Select Contig

Create Target

The wizard will allow the creation of a new Target. The other ways to create a target are to use the 'Import Target' wizard (and import annotations plus create a Target in a single step), or the 'Create Profile' wizard, which merges the 'Import Reference Data', 'Import Known Variants' and 'Import Target' wizards into one.

A Target is defined as a list of annotations. These annotations can be for genes, exons or amplicons. The annotations belong to a particular Reference Genome and can be imported into the tool via the Reference Genome Dashboard.

List of wizard steps:

- Create Target

It is not sufficient to simply create a Target - once a Target has been created it should be configured by clicking on it (which will navigate to the Target Dashboard), and then configuring the Target via the Configure Target wizard.

Configure Target

This function can be used to add and remove annotations from a Target. The annotations will already need to exist - either imported against a Reference, from the References screen, or by using the 'Import Target Annotations' function in the Targets screen, or the 'Create Profile' function in the main Data Analysis screen.

Import Target Annotations

This function will import annotations for genes, exons or amplicons and use them to create a Target.

The reference genome will need to be created first (if it doesn't exist already), and all the reference data should already be imported within Omixon Target.

The easier alternative to this function is to use the 'Create Profile' wizard, which merges the 'Import Reference Data', 'Import Known Variants' and this 'Import Target' wizard into one function.

Export Mapped Data

Any mapped short read data that has been mapped within Omixon Target (or imported using the 'Import Mapped Data' function) can be

exported again, either in .sam or .bam format.

The export will also create a standard index file for the exported data.

Export Mapped Regions

From the Genome Browser, it is possible export the short read that overlap or are contained within the currently visible region.

There are two options for the export - the reads can be exported in either .fastq format (for re-mapping) or in .sam/.bam format.

Export Approved Variants

This function is currently the final step in the 'data analysis' pipeline. Once you have a short-list of approved variants you can export an annotated .vcf file from the tool using this function. Only variants that have been 'approved' (either by the manual or auto approval functions) can be exported in this way.

Export Activity Log

This function will export the activity log into .txt, .csv or .xls format. The log can be for a single sample, a whole analysis, a whole experiment or the whole application.

It is useful for obtaining a history of everything that was done within the tool for that sample, for the purposes of documenting an experiment.

Reset Everything

It is possible to Reset everything - reset the whole of Omixon Target - to an empty state.

Resetting will delete all imported and analysed data.

Resetting is only recommended if it's the first thing that is done with the tool, if the tool is being used for demonstration purposes, or if you really do want to start again completely from scratch.

Reset Everything is available from the Settings Dashboard. You will be asked again if you are completely sure before the reset process is started. There is no way to recover data that is removed by a reset.

Once you have reset, the tool will then only contain the Expert Profile. The next recommended step is to use 'Add Profile' to add one (or more) of the pre-configured Profiles to the tool, which will download and import all the data required for each Profile.

Appendix

Keyboard shortcuts

The following generic keyboard shortcuts are available in Omixon Target.

- **F8** key closes the window and exits the application
- **F11** key switches to fullscreen display mode to utilize as much space as possible for visualization - note that certain platforms and window managers are not working properly together with this function, the first time you use you will get a warning message about this