

Instructions for Use



LymphoTrack® Dx Software – MiSeq™ Version 2.4.3

This software package is provided with each LymphoTrack Dx Assay for the Illumina MiSeq instrument to analyze raw FASTQ files. The software is composed of three distinct parts: the LymphoTrack Data Analysis Application, a LymphoTrack Reporter plugin that runs automatically after analysis, and an optional LymphoTrack Visualization method consisting of four Microsoft Excel®.xslm spreadsheets. The software runs on supported Microsoft Windows® platforms.

This software package allows for clonality analysis of single or multiple target data (*IGHV* Leader (or “Leader”), *IGH* FR1, *IGH* FR2, *IGH* FR3, *IGK*, *TRB* or *TRG*) and additionally provides the somatic hypermutation status of samples run with the Leader or *IGH* FR1 assays.

IVD For *In Vitro* Diagnostic Use.

Product Compatibility

This LymphoTrack Dx Software – MiSeq (**REF** 95000009) was developed specifically for and is compatible with **only** the following Invivoscribe assays:

Catalog #	Description	UDI
REF 91210059	LymphoTrack Dx <i>IGHV</i> Leader Somatic Hypermutation Assay Kit A – MiSeq	00850052003814
REF 91210069	LymphoTrack Dx <i>IGHV</i> Leader Somatic Hypermutation Assay Panel – MiSeq	00850052003821
REF 91210009	LymphoTrack Dx <i>IGH</i> FR1 Assay Kit A – MiSeq	00850052003647
REF 91210039	LymphoTrack Dx <i>IGH</i> FR1 Assay Panel – MiSeq	00850052003654
REF 91210089	LymphoTrack Dx <i>IGH</i> FR2 Assay Kit A – MiSeq	00850052003838
REF 91210099	LymphoTrack Dx <i>IGH</i> FR2 Assay Panel – MiSeq	00850052003845
REF 91210109	LymphoTrack Dx <i>IGH</i> FR3 Assay Kit A – MiSeq	00850052003852
REF 91210119	LymphoTrack Dx <i>IGH</i> FR3 Assay Panel – MiSeq	00850052003869
REF 91210129	LymphoTrack Dx <i>IGH</i> FR1/2/3 Assay Kit A – MiSeq	00850052003876
REF 91210139	LymphoTrack Dx <i>IGH</i> FR1/2/3 Assay Panel – MiSeq	00850052003883
REF 91220009	LymphoTrack Dx <i>IGK</i> Assay Kit A – MiSeq	00850052003906
REF 91220019	LymphoTrack Dx <i>IGK</i> Assay Panel – MiSeq	00850052003913
REF 92270019	LymphoTrack Dx <i>TRG</i> Kit A – MiSeq	00850052003678
REF 92270009	LymphoTrack Dx <i>TRG</i> Panel – MiSeq	00850052003661
REF 92250009	LymphoTrack Dx <i>TRB</i> Kit A – MiSeq	00850052003937
REF 92250019	LymphoTrack Dx <i>TRB</i> Panel – MiSeq	00850052003944

Intended Use

The LymphoTrack Dx Software – MiSeq package is intended to be used for bioinformatics data analysis and visualization of raw FASTQ files from the Illumina MiSeq instrument when used with LymphoTrack Dx Assays.

Minimum System Requirements

- Processor: Intel Core 2 Duo or newer CPU recommended.
- Hard Drive: At least 80 GB of free disk space is required; 250 GB recommended.
- RAM: 4 GB required; 8 GB or more recommended.
- Operating System: Windows 10 (64-bit) is required.
- Java 8 for 64-bit operating systems (Java configured for 32-bit operating systems is not compatible with the LymphoTrack Dx Software – MiSeq). The most recent version of Java can be downloaded directly from Oracle at <http://www.java.com/>.
- A PDF reader, such as Adobe Acrobat Reader, to visualize data reports generated by the LymphoTrack Reporter.
- Microsoft Office Excel 2007, 2010 or 2013 with Macro support enabled is required for data visualization.
- A CD-ROM drive.
- The LymphoTrack Dx Software – MiSeq v2.4.3 has been validated for use with Windows 10, configured for English (United States) language and default display size settings, and Microsoft Excel 2007/2010/2013 for alternate data visualization. Use of other Windows/Excel versions and/or language/display size settings may not be compatible.

Warnings and Precautions

- **Default screen size.** Optimal screen resolution for LymphoTrack Dx Software – MiSeq is 100%. Increasing the monitor default screen size above 100% may affect the display of field text in the user interface.
- **System font.** User interface is designed to use the default system font settings on a Windows computer using the English (US) language locale. Changes to the system font settings or using a language locale other than English (US) may impact the display of field text in the user interface.
- **Compatible files.** LymphoTrack Dx Software – MiSeq is compatible with *fastq.gz* files generated by the MiSeq Reporter Software. LymphoTrack Dx Software – MiSeq is not compatible with *fastq.gz* files generated by the *Bcl2fastq* software.
- **Adapter trimming.** Adapter trimming must be selected ☒ when creating MiSeq sample sheets so the resulting *fastq.gz* files generated are adapter trimmed.
- **Sleep or hibernate settings.** Consider disabling the sleep and/or sleep or hibernate mode features before executing the LymphoTrack Dx Data Analysis Application. If the computer enters sleep or hibernate mode, the software may stop running.
- **Characters in pathname and file name.** 1) Avoid spaces in the pathname for the data files or software (pathnames include file folders and file names) and more than one consecutive space is not permitted. 2) It is important that the filenames only contain the following characters (A-Z, a-z, 0-9, ., - (hyphen)). If the software encounters a character not within this set, it may fail. Please make sure to follow this sample naming nomenclature when setting up the Sample Sheet for the MiSeq. 3) Do not exceed 255 characters for a pathname. Sample names that cause the output filenames to approach this limit are difficult to copy to other locations.
- **Duplicate samples must have a unique sample name.** Give each sample a unique name and identifier when naming samples. If samples are run in duplicate, they can be named “Sample1a” and “Sample1b”. Failure to provide unique names to samples that will be run together on the same flow cell will result in only one sample being analyzed by the software during the analysis process.

1. Installing the Software

- 1.1. Copy the *LymphoTrackDxMiSeq2.4.3.zip* archive from the software CD to a local drive on the computer. Install software only on a local drive and not a network drive. Running the software across a network connection may not function properly.
- 1.2. Extract the archive file that was copied to the local computer in the previous step. This can typically be accomplished by right clicking on the file and choosing *Extract All* from the contextual menu.

A folder named *LymphoTrackDxMiSeq2.4.3* is created once the archive is extracted.

2. Running the Data Analysis Software

- 2.1. Identify the paired-end, adapter-trimmed **.fastq.gz* files to be analyzed. These files are located within the MiSeq run folder and are the **.fastq.gz* files that will be used as input for the **LymphoTrackDxMiSeq** data analysis application. The *Undetermined_*.fastq.gz* files are not required for analysis.
- 2.2. Copy the **.fastq.gz* files identified above to an empty folder on the computer. Download these files to the same local computer which has the LymphoTrack Dx Software installed. The drive containing this folder requires at least 80 GB of available space. This folder will be the **<input_folder>** referenced below for the software.
- 2.3. Double-click on the **startLymphoTrackDxMiSeq** file located in the *LymphoTrackDxMiSeq2.4.3* folder copied to the computer during installation.
- 2.4. Read the license agreement. To accept the terms, click the **Accept** button to proceed.
- 2.5. Check the desired analysis mode from the *Choose analysis mode(s)* checkbox options. The analysis mode must correspond to the type of data in the run (Leader, *IGH FR1*, *IGH FR2*, *IGH FR3*, *IGK*, *TRB* and/or *TRG*).
- 2.6. Select the desired decimal mark notation from the *Choose decimal mark options*. This option impacts the use of the Microsoft Excel data visualization macro later, and is a required input. The choices are to use a decimal point (".") or a decimal comma (",") for the decimal mark.

IMPORTANT!

Decimal mark. To determine which option to choose, open a new Excel worksheet, and type **=10000** into cell A1. Right-click this cell, and select **Format Cells**.

- ☐ From the *Number* tab, look in the *Category:* list, and select **Number**. To the right, the *Decimal places* value which is set to **2**.
- ☐ Below this value is a *Use 1000 Separator* checkbox. Check this box, and look at the sample number above. It will contain the original number entered, formatted according to the Excel defaults for a number with a decimal mark.
- ☐ If this number appears as '*10,000.00*', then select the **decimal point** option. If this number appears as '*10.000,00*', then select the **decimal comma** option.

- 2.7. Click the **Browse** button to choose an input folder. This folder is the **<input_folder>** referenced above and contains the files to be analyzed. When the input folder is selected it will appear empty.

IMPORTANT!

The <input_folder> selected must not contain any sample files with the same name as the input folder. If the program was run with a sample that shared the same filename as the folder that was provided as input to the analysis functionality, the program will freeze and analysis runtime will increase significantly (*i.e.*, it will take 30 minutes to process a small sample that takes seconds to process normally).

- 2.8. Once the **<input_folder>** and analysis mode are selected, click the **Launch Program** button.

IMPORTANT!

The <input_folder> selected must contain only *.fastq or *.fastq.gz files. If neither file type nor a mix of file types are present, the program will display an error message and request a different folder to be chosen. If the **.fastq* files are generated by *Bcl2fastq* (the filenames contain the barcode sequence instead of the "S1, S2, S3" etc... nomenclature), the program will display an error message and request the data to be re-analyzed using the MiSeq Reporter Software.

- 2.9. Status of analysis.** After clicking the **Launch Program** button, the main program window will display the program's status. When the program is finished analyzing the input data files, the main program window will display the message *Analysis is complete...* and show the location of the output folder as well as provide a prompt to open the appropriate Excel Data Visualization spreadsheet. A new analysis can be started from the main program window as soon as the previous analysis is complete.

IMPORTANT!

Length of analysis. The size and number of samples analyzed, as well as the analysis mode and the computer's hardware will determine the amount of time the program takes to run. A typical read distribution for a 24-sample run is 200,000 to 1 million reads per sample. These numbers may change with future releases of the MiSeq platform.

Analysis of a typical single *IGHV* Leader, *IGH* FR1, *IGH* FR2, *IGH* FR3, *IGK*, *TRB*, or *TRG* sample requires approximately 1-2 minutes to complete. A typical 24-sample Leader-only, *IGH* FR1-only, *IGH* FR2-only, *IGH* FR3-only, *IGK*-only, *TRB* only or *TRG*-only run requires approximately 25-50 minutes to complete. Analysis of a typical 24-sample combined Leader, *IGH* FR1, *IGH* FR2, *IGH* FR3, *IGK*, *TRB* and/or *TRG* run multiplies the analysis time by the number of targets being analyzed.

- 2.10. Output folder(s).** Depending on the analysis mode chosen, one or more output folders will be created within the **<input_folder>** containing the source **.fastq* or **.fastq.gz* files. Table 1 outlines the relationship between the analysis mode and the name of the output folder(s).

Table 1. Output folder name according to analysis mode

Analysis Mode	Output folder name(s)
Leader	<i>LEADER_output</i>
<i>IGH</i> FR1	<i>IGH_FR1_output</i>
<i>IGH</i> FR2	<i>IGH_FR2_output</i>
<i>IGH</i> FR3	<i>IGH_FR3_output</i>
<i>TRB</i>	<i>TRB_output</i>
<i>TRG</i>	<i>TRG_output</i>
<i>IGK</i>	<i>IGK_output</i>
Any combination of: Leader, <i>IGH</i> FR1, <i>IGH</i> FR2, <i>IGH</i> FR3, <i>IGK</i> , <i>TRB</i> and <i>TRG</i>	Two to seven output folders listed above

- 2.11.** After the program has finished running successfully, the output folder will contain several TSV files and a PDF file for each sample processed. Note that "*" in the tables below represents the sample name as entered into the sample sheet, followed by platform specific information (e.g. *samplename_S1_L001_001_combined*).
- 2.12.** For Leader, *IGH* FR1, *IGH* FR2, *IGH* FR3, *TRB* and *IGK* analysis; please see Table 2 for descriptions of the nine files provided for each sample.

Table 2. Description of *IGH*, *TRB* and *IGK* output files.

Leader, <i>IGH</i> FR1, <i>IGH</i> FR2, <i>IGH</i> FR3, <i>TRB</i> and <i>IGK</i> Analysis	
<i>*.fastq_read_summary.tsv</i>	<p>Contains a list of the top 200 unique sequences sorted by count (number of reads), specifying:</p> <ul style="list-style-type: none"> Which V-J recombination event it best fits Length Actual % of total reads Cumulative % of total reads How well the read covers the identified V-gene targeted by the primers (%) (<i>IGH/IGK</i> and <i>TRB</i>) Putative CDR3 sequence <p>Leader and <i>IGH</i> FR1 also contains:</p> <ul style="list-style-type: none"> Mutation rate to partial V-gene (%) <p>Leader, <i>IGH</i> FR1, <i>IGH</i> FR2, <i>IGH</i> FR3 also contains:</p> <ul style="list-style-type: none"> If the rearrangement is predicted to be in-frame (Y/N/n/a). If either the V or J gene segment usage cannot be predicted due to the quality of the alignment, the software will not be able to predict if the gene is in-frame and the result will be n/a. If the rearrangement is predicted to not contain a stop codon (Y/N). When an in-frame prediction cannot be made a prediction for whether a stop codon is present cannot be accurately made.
<i>*.fastq_read_summary_merged_top10_searchtop500.tsv</i>	Same format as the <i>read_summary</i> file, but only includes the top 10 reads from the <i>read_summary</i> merged with reads in the top 500. A read will merge with another if they are only 1 or 2 nucleotides (nts) different, with the read with the highest count maintaining its sequence identity, and the read with the lower count added to the higher one.
<i>*.fastq_unique_reads.tsv</i>	A <i>*.fasta</i> file of unique sequences, sorted by count.
<i>*.fastq_CDR3_unique_reads.tsv</i>	A <i>*.fasta</i> file of putative CDR3 sequences found in the ' <i>*.fastq_unique_reads</i> ' file.
<i>*.fastq_VJ_sequence_frequency_family.tsv</i>	Each family is a grouping of all possible recombinations found in a particular V or J region such as a V ₁ family or a J ₄ family. Within the V ₁ family any number of recombinations can be found such as V ₁₋₆₉ and V ₁₋₈ . This file contains a list of all possible V family – J family recombinations, each with a sorted list of reads (from highest to lowest %). This file is best used in a “stack” graph in Excel.
<i>*.fastq_read_summary_family.tsv</i>	<p>Contains a sorted list of unique sequences, by count, specifying:</p> <ul style="list-style-type: none"> Which V-J family recombination event it best fits Length Actual % of total reads Cumulative % of total reads How well the read covers the identified V-gene targeted by the primers (%) (<i>IGH/IGK</i>) Putative CDR3 sequence <p>Leader and <i>IGH</i> FR1 also contains:</p> <ul style="list-style-type: none"> Mutation rate to partial V-gene (%) <p>Leader, <i>IGH</i> FR1, <i>IGH</i> FR2, <i>IGH</i> FR3 also contains:</p> <ul style="list-style-type: none"> If the rearrangement is predicted to be in-frame (Y/N/n/a) If the rearrangement is predicted to not contain a stop codon (Y/N)
<i>*.fastq_VJ_usage.tsv</i>	A count matrix and % matrix of the different V-J recombination events.
<i>*.fastq_VJ_usage_family.tsv</i>	A count matrix and % matrix of the different V family – J family recombination events.
<i>*<timestamp>.pdf</i>	The <i>*.pdf</i> report generated by the LymphoTrack Reporter will contain similar information as the “merged read summary” file, the “read summary” file, a ‘V – J’ usage’ graph, and a ‘V – J sequence frequency’ graph. See section 3 <i>Data Visualization Using LymphoTrack Reporter (Automated, Recommended)</i> for more information.

Note: *IGK* analysis additionally includes information for other elements in the *IGK* locus involved in recombinations (IGKINTR and IGKDEL).

Note: *TRB* analysis additionally includes information for other elements in the *TRB* locus involved in recombinations of D – J region.

Note: ‘Putative CDR3 sequence’ is determined by matching an exact definition of a CDR3 sequence. Any mutation of anchor points or disruption of the reading frame will result in a value of ‘not found’.

□ For TRG analysis, please see Table 3 for descriptions of the seven files provided for each sample.

Table 3. Description of TRG output files.

TRG Analysis	
<i>*.fastq_read_summary.tsv</i>	A list of unique sequences sorted by count, specifying: <ul style="list-style-type: none"> ▪ Which V-J recombination event it best fits ▪ Length ▪ Actual % of total reads ▪ Cumulative % of total reads ▪ Putative CDR3 sequence
<i>*.fastq_read_summary_merged_top10_searchtop500.tsv</i>	Same format as the <i>read_summary</i> file, but only includes the top 10 reads from the <i>read_summary</i> merged with reads in the top 500. A read will merge with another if they are only 1 or 2 nts different, with the read with the higher count maintaining its sequence identity, and the read with the lower count added to the higher one.
<i>*.fastq_unique_reads.tsv</i>	A <i>*.fasta</i> file of unique sequences, sorted by count.
<i>*.fastq_CDR3_unique_reads.tsv</i>	A <i>*.fasta</i> file of putative CDR3 sequences found in the <i>*.fastq_unique_reads</i> file.
<i>*.fastq_VJ_sequence_frequency.tsv</i>	A list of all possible V-J recombinations, each with a sorted list of reads (from highest to lowest %). The first number represents the % of total reads that the first unique read accounted for out of all reads sequenced. This file is best used in a "stack" graph in Excel.
<i>*.fastq_VJ_usage.tsv</i>	A count matrix and % matrix of the different V-J recombination events.
<i>*<timestamp>.pdf</i>	The <i>*.pdf</i> report generated by the LymphoTrack Reporter will contain similar information as the "merged read summary" file, the "read summary" file, a "VJ usage" graph, and a "VJ sequence frequency" graph. See section 3 <i>Data Visualization Using LymphoTrack Reporter (Automated, Recommended)</i> for more information.

Note: 'Putative CDR3 sequence' is determined by matching an exact definition of a CDR3 sequence. Any mutation of anchor points or disruption of the reading frame will result in a value of "not found".

2.13. The output folder(s) contain **.pdf* reports and **.tsv* output data files. The **.tsv* files can be used as inputs for the *Data Visualization* spreadsheet(s).

3. Data Visualization Using LymphoTrack Reporter (Automated, Recommended)

- 3.1.** A *.pdf report for each sample is automatically generated at the completion of analysis. The report is named after the sample name present in the sample sheet, followed by platform specific information, and ends with a timestamp (DDMMYYhhmmss).

IMPORTANT! **Caution when saving the *.pdf report:** Depending on the program used to open the *.pdf, one or more of the columns can be edited. Do NOT make changes to any part of the *.pdf. If the user suspects an edit has been made, do NOT save the *.pdf. If the user suspects an edit has been saved, re-run the analysis to generate a new *.pdf report.

- 3.2.** The first page of the *.pdf report will contain an **Index Q30** value. This value is calculated as the percentage of bases that have a quality score of 30 or higher, but only from reads that are assigned to an index after trimming. It can also be used to determine if the index passed the standard Illumina specification. In some cases, the run Q30 score provided by the sequencer will underestimate the quality of a run because of counting undetermined sequences and untrimmed regions towards the Q30. Undetermined and untrimmed regions are typically lower quality and do not contribute to the downstream analysis by LymphoTrack.
- 3.3.** The first page of the *.pdf report will contain a table of the **merged read summary** results, followed by the graphs of the **V – J Sequence Frequency** and **V – J Usage** results, and the remaining pages contain a table of the unmerged **read summary** results. The contents of these are detailed in Table 4:

Table 4. Description of PDF report contents.

PDF Output for LymphoTrack Reporter	
<i>Read Summary</i>	<p>Contains a sorted list of the top 200 unique sequences by count, specifying:</p> <ul style="list-style-type: none"> Which V – J recombination event each sequence best fits <ul style="list-style-type: none"> Sequence Sequence length Actual % of total reads Cumulative % of total reads listed for top 200 reads <ul style="list-style-type: none"> How well the read covers the identified V-gene targeted by the primers (%). (IGH/IGK) Putative CDR3 sequence <p>Leader and IGH FR1 also contains:</p> <ul style="list-style-type: none"> Mutation rate to partial V-gene (%) <p>Leader, IGH FR1, IGH FR2, IGH FR3 also contains:</p> <ul style="list-style-type: none"> If the rearrangement is predicted to be in-frame (Y/N/n/a). If either the V or J gene segment usage cannot be predicted due to the quality of the alignment, the software will not be able to predict if the gene is in-frame and the result will be n/a. If the rearrangement is predicted to not contain a stop codon (Y/N). When an in-frame prediction cannot be made a prediction for whether a stop codon is present cannot be accurately made.
<i>Merged Read Summary</i>	Same format as the Read Summary worksheet, but only for the top 10 reads after merging with the top 500 if reads are only 1 or 2 nts different from each other. Additionally, the index Q30 score is displayed under the Total Read Count.
<i>V – J Sequence Frequency Graph</i>	A histogram showing the top 200 sequence frequencies of each V family – J family (for IGH/IGK and TRB) or V gene – J gene (for TRG) recombination.
<i>V – J Usage Graph</i>	A 3D bar chart showing the V family-J family usage of the top 200 sequences (for IGH/IGK and TRB) or V gene – J gene usage for all sequences (for TRG).

Note: Putative CDR3 sequence is determined by matching an exact definition of a CDR3 sequence. Any mutation of anchor points or disruption of the reading frame will result in a value of “not found”.

Note: IGK analysis additionally includes information for other elements in the IGK locus involved in recombinations (IGKINTR and IGKDEL).

Note: TRB analysis additionally includes information for other elements in the TRB locus involved in recombinations of D – J region.

4. Data Visualization (Manual): Running the Excel Data Visualization Spreadsheets - Discommended

- 4.1.** Check the Security settings within Microsoft Excel to make sure Macro support is properly enabled:
- ☐ Excel 2007: Click the **Office** button and choose **Excel Options**.
 - ☐ Excel 2010/2013: Click the **File** menu and choose **Options**.
 - ☐ From the *Options* window, click on **Trust Center** and then the **Trust Center Settings...** button.
 - ☐ Click on **Macro Settings** and set the appropriate Macro security level. The *Disable all macros with notification* option is the recommended setting.
 - ☐ Click **Ok** and **Ok** again.
 - ☐ Close Excel.
- 4.2.** Open the appropriate Data Visualization spreadsheet file by double-clicking it. The visualization files are located in the **LymphoTrackDxMiSeq_2.4.3** folder previously copied from the software CD to the computer:
- ☐ *LymphoTrackDxIGHVisualization.xlsm* for Leader, IGH FR1, IGH FR2, IGH FR3
 - ☐ *LymphoTrackDxIGKVisualization.xlsm* for IGK
 - ☐ *LymphoTrackDxTRBVisualization.xlsm* for TRB
 - ☐ *LymphoTrackDxTRGVisualization.xlsm* for TRG
- 4.3.** When opening the file, the message: **Security Warning - Macros have been disabled** may appear.
- ☐ Excel 2007: Click the **Options...** button to the right of the warning and select **Enable this content** and click **Ok**.
 - ☐ Excel 2010/2013: Click the **Enable Content** button.
- The data visualization macro will start automatically.
- 4.4.** The License Agreement will be displayed. The license agreement must be accepted to use the software.
- 4.5.** After accepting the License Agreement, a dialog box will open with the prompt: *Please select the Read Summary file for the sample you want.* The macro can only perform data visualization on one sample at a time, so this procedure will need to be repeated for each sample to be visualized.
- 4.6.** Browse to the output folder created by the Data Analysis application that contains the data to be visualized, and select the **.fastq_read_summary* file for the sample to be examined.
- 4.7.** Click the **Open** button.
- 4.8.** The macro will load the data files for the chosen sample and create several worksheets, each of which will present different readable views of the data. This may take a few moments, depending on the size of the sample and the computer's hardware.

IMPORTANT!

Due to a limitation in Microsoft Excel, the visualization software is not capable of displaying more than 1,048,576 rows of data per sheet. Only the first 1,048,576 rows of data will be shown if any of the Data Analysis output files for the current sample contain more rows than this limit.

If this limit is reached, the program will display a message stating which sheet(s) are affected by this limitation. Click **OK** to proceed.

Though not all data may be displayed, the graphs are not affected and will still be accurate.

IMPORTANT!

Graphs are generated assuming a **Left-to-right** display direction. In countries where the normal reading direction is right-to-left, the default Excel settings may be incompatible and graphs are not generated correctly. To fix this, the Excel display option must be changed to **Left-to-right**. This option can be set from either *Office > Excel Options* in Excel 2007, or *File > Options* in Excel 2010/2013. Select **Advanced**, then scroll down to the *Display* settings. Change the *Default direction* option to **Left-to-right**, then click **OK**. The next time Excel is opened, the worksheets will read from left-to-right, and the visualization software will generate charts correctly.

If this situation occurs and further assistance is required, please contact Invivoscribe Technical Support department at +1 (858) 224-6600, Monday through Friday 7:00AM – 5:00PM PST/PDT, or via email at support@invivoscribe.com for assistance.

4.9. If a message regarding *Complex Formatting* appears, click **Yes**.

4.10. **Saving the visualization file.** When processing is complete, a message will appear stating the program has completed. Saving the file at this point is strongly recommended.

IMPORTANT!

When saving the file, make sure to choose **Save As** and choose an Excel Workbook (*.xlsx or *.xls), because the default behavior of choosing **Save** or clicking the disk button will be to replace the *.xsm Macro with the current workbook. This will require reinstallation of the macro with a new copy from the software CD for subsequent analyses.

4.11. If a warning message appears regarding saving as a macro-free workbook, click the **Yes** button.

4.12. The completed workbook for Leader, *IGH* FR1, *IGH* FR2, *IGH* FR3, *IGK*, *TRB* and *TRG* will include the following eight worksheets described in Table 5.

Table 5. Worksheets included Excel output file.

Completed Workbook for Leader, <i>IGH</i> FR1, <i>IGH</i> FR2, <i>IGH</i> FR3, <i>IGK</i> , <i>TRB</i> and <i>TRG</i>	
<i>Read Summary</i>	<p>Contains a sorted list of the top 200 unique sequences by count, specifying:</p> <ul style="list-style-type: none"> Which V-J recombination event each sequence best fits Sequence Sequence length Actual % of total reads Cumulative % of total reads listed for all reads (<i>TRG</i> only) Cumulative % of total reads listed for top 200 reads (<i>IGH/IGK</i> and <i>TRB</i>) How well the read covers the identified V-gene targeted by the primers (%) (<i>IGH/IGK</i>) Putative CDR3 sequence <p>Leader and <i>IGH</i> FR1 also contains:</p> <ul style="list-style-type: none"> Mutation rate to partial V-gene (%) <p>Leader, <i>IGH</i> FR1, <i>IGH</i> FR2, <i>IGH</i> FR3 also contains:</p> <ul style="list-style-type: none"> If the rearrangement is predicted to be in-frame (Y/N/n/a). If either the V or J gene segment usage cannot be predicted due to the quality of the alignment, the software will not be able to predict if the gene is in-frame and the result will be n/a. If the rearrangement is predicted to not contain a stop codon (Y/N). When an in-frame prediction cannot be made a prediction for whether a stop codon is present cannot be accurately made. Putative CDR3 sequence.
<i>Merged Read Summary</i>	Same format as the <i>Read Summary</i> worksheet, but only for the top 10 reads after merging with the top 500 if reads are only 1 or 2 nts different from each other.
<i>VJ Sequence Frequency Graph</i>	A histogram showing the top 200 sequence frequencies of each V family – J family (for <i>IGH/IGK</i> and <i>TRB</i>) or V gene – J gene (for <i>TRG</i>) recombination.
<i>VJ Usage</i>	Raw counts and frequencies of the top 200 V family - J family (for <i>IGH/IGK</i> and <i>TRB</i>) or V gene – J gene (for <i>TRG</i>) recombinations.
<i>VJ Usage Percent Graph</i>	A 3-D graph showing the top 200 sequence frequencies of each V family – J family (for <i>IGH/IGK</i> and <i>TRB</i>) or V gene – J gene (for <i>TRG</i>) recombination.
<i>VJ Usage Raw Graph</i>	A 3-D graph showing the top 200 raw counts of each V family – J family (for <i>IGH/IGK</i> and <i>TRB</i>) or V gene – J gene (for <i>TRG</i>) recombination.
<i>VJ Sequence Frequency</i>	A “stack” list of all possible V family – J family (for <i>IGH/IGK</i> and <i>TRB</i>) or V gene – J gene (for <i>TRG</i>) recombinations, each with a sorted list of the top 200 reads (from highest to lowest %).
<i>Unique Reads</i>	A *.fasta file of unique sequences, sorted by count.

Note: *IGH* listed in the table above refers to Leader, *IGH* FR1, *IGH* FR2, and *IGH* FR3 analysis. The frequency graphs for these gene targets only specify *IGH* in the title. The data output folder specifies whether Leader or one of the framework (FR) regions was analyzed by the software and this information is included in the data file location field at the top of the Read Summary worksheet.

Note: *IGK* analysis additionally includes information for other elements in the *IGK* locus involved in recombinations (IGKINTR and IGKDEL).

Note: *TRB* analysis additionally includes information for other elements in the *TRB* locus involved in recombinations of the D – J region

Note: Top 200: When specified, the worksheets will show the top sequences (up to 200) for Leader, *IGH* FR1, *IGH* FR2, *IGH* FR3, *TRB* and *IGK*. All of the sequences are shown for *TRG* in the Read Summary.

Note: ‘Putative CDR3 sequence’ is determined by matching an exact definition of a CDR3 sequence. Any mutation of anchor points or disruption of the reading frame will result in a value of ‘not found’.

IMPORTANT!

V/J gene assignments found in the *Read Summary* can sometimes be 'none'. These assignments are based on the quality of the best alignment between the read and the reference genes. If the best alignment falls below a quality threshold, the gene assignment cannot be confidently made, and assigned 'none'.

IMPORTANT!

The Mutation rate to partial V-gene (%) will display 'n/a' for *IGH* FR2 and FR3. There is limited V-gene coverage due to the shorter amplicon sizes of these targets which results in inaccurate mutation rate calculations.

- When analyzing the somatic hypermutation status of samples, focus on the top 4 reads (limit monoclonal or biclonal specimens to the top 2 or 4 reads, respectively), as the somatic hypermutation rate (SHM) and in-frame/stop-codon analyses are most informative for these reads. Due to the unbiased nature of the NGS platforms, many more reads than just the top 4 are sequenced and analyzed, most of which are low abundance rearrangements, or can be attributed to artifacts in amplification or sequencing. SHM and in-frame/stop-codon analyses are not useful in non-clonal reads in these cases. Please note that Invivoscribe uses GRCh37 for reference sequences and results may differ slightly from the IMGT database. In addition, many pseudogenes and non-functional genes are not included in the analysis. Interpret rearrangements with somatic hypermutation statuses at or near the 2.0% cutoff with caution, particularly when running the *IGH* FR1 assay which only targets a portion of the *IGHV* region as the sequence upstream of the primer binding site will not be assessed.

5. Technical and Customer Service

We appreciate your business. We are happy to assist you with understanding this software, and will provide ongoing technical assistance Monday through Friday to keep the assays performing efficiently in your laboratory.

Contact Information



Invivoscribe, Inc
















10222 Barnes Canyon Road | Building 1 | San Diego | California 92121-2711 | USA

Phone: +1 858 224-6600 | Fax: +1 858 224-6601 | Business Hours: 7:00AM - 5:00 PM PST/PDT

Technical Service: support@invivoscribe.com | Customer Service: sales@invivoscribe.com | Website: www.invivoscribe.com

6. Symbols

The following symbols are used in Invivoscribe NGS product labeling.

	Catalog Number		Expiration Date
	Reagent Volume		Authorized Representative in the European Community
	Lot Number		Consult Instructions for Use
	Storage Conditions		For <i>In Vitro</i> Diagnostic Use
	Unique Device Identifier		Manufacturer
	UK Conformity Assessed		UK Responsible Person
	Swiss Authorized Representative		European Conformity
	Software Application		

7. Software Frequently Asked Questions

Q1: When **'Launch Program'** is clicked, the following (or similar) error occurs:

Application terminated with error: Cannot run program ".\resources\Parse_TRG_miseq.exe": CreateProcess error=2, the system cannot find the file specified.

For help with this error, please contact Invivoscribe Technical Support.

A1: Was the compressed program file unzipped (see section 1 *Installing the Software*)? This error commonly occurs when trying to run the program before decompressing the program file. Windows allows the ability to view the contents of a zip file simply by double-clicking the file, and the contents appear to be complete and functional, but it will not run properly from this state.

Q2: Why does the Read Summary from the *IGH*, *IGK* or *TRB* analysis have fewer than 200 reads?

A2: The LymphoTrack Dx Software - MiSeq aligns the top 200 reads during analysis, but the *Read Summary* only reports reads that have aligned to a V or J gene found in the reference. If a read cannot align to the reference, it is left out of the Read Summary.

The reasons for a non-alignment to the reference are numerous, but usually focus on a problem with the sample, leading to a large amount of non-specific amplification to occur. This event can stem from low numbers of target cells in the sample, improper PCR conditions, or low-quality sample types, such as fragmented FFPE. To reduce the chance of these events, please follow the Instructions for Use precisely.

Q3: Why does the V-assignment or J-assignment say **'none'** in the Read Summary table?

A3: To assign a read to a specific V and/or J gene, a minimum quality threshold must be attained during the alignment. If the alignment is poor, a confident assignment cannot be made and is instead labeled **'none'**. Note that a read cannot be assigned **'none'** for both V and J genes, as these would not be reported, per the condition discussed in question 2.

Q4: The **'V-J Sequence Frequency Graph'** has strange values for the y-axis. The **'% total reads'** is in the millions, how is this possible?

A4: This is likely due to running the LymphoTrack Dx program using the incorrect decimal type for your region. See section 2 *Running the Data Analysis Software*, specifically for selecting the decimal mark.

Q5: All of the graph worksheets are empty, but the non-graph worksheets have data. Where did the graphs go?

A5: This may be due to the default direction of the Excel program being *right-to-left* instead of *left-to-right*. If the column labels on an empty worksheet start at **A** on the right-side of the worksheet, the direction will need to be switched to *left-to-right* before using the data visualization macro. To change this setting, go to *Excel options* from the **Office** button in the upper left-hand corner, click on **Advanced** and scroll down to the *Display* section. An option called *Default direction* is listed here and selecting **Left-to-right** will allow subsequent data visualization runs to work correctly. Note: This will permanently change the direction until it is changed back. Closing Excel will not revert this change.

Q6: The program stopped after clicking **'Launch Program'** and produced this error:

Application terminated with error: Filename not compatible. Please use only MiSeq MSR software generated FASTQ files.

For help with this error, please contact Invivoscribe Technical Support.

A6: The LymphoTrack Dx software – MiSeq expects paired-end *.fastq files generated by the MiSeq Reporter (MSR) software contained within the MiSeq machine. If the MSR software was not used to generate sequence files, the base-calling must be performed using the MSR software. Please refer to the MiSeq User Guide for more information.

Performing the base-calling using the MSR software is important because, while other programs can generate very similar results as the MSR software (for example, Bcl2fastq), these programs can also generate very different results, depending on the specific parameters. The LymphoTrack Dx Assay is only valid when analyzing *.fastq files generated from the MSR software.

Q7: Why do the analysis results seem to include fewer reads than expected?

A7: If there are very few or zero total reads analyzed according to the Read Summary, make sure the correct analysis type is selected (Leader, IGH FR1, IGH FR2, IGH FR3, IGK, TRB or TRG) when running the LymphoTrack Dx software – MiSeq. If there seems to be a depressed number of total reads, compare the size of the *.fastq files being analyzed with size of the FASTQ files present on the MiSeq instrument. If the MiSeq instrument has larger files, then r-transfer the sequence data to ensure analysis of the full dataset.

The MiSeq sequencer displays different messages when processing the sequencing data. One message may state ‘Sequencing Complete’, and another may later state ‘Analysis Complete’. Transferring the FASTQ data in between these two messages may result in an incomplete LymphoTrack Dx analysis.

Q8: Why does the Read Summary contain a read in the top ten that is not in the Merged Read Summary?

A8: The “disappearing” read may have been merged with a read ranked above it, or it is possible the alignment of this read was exceptionally poor. During the merging process, reads in the read summary are aligned with those in the unique reads file; however, reads in the read summary have been oriented to match the alignment to the reference, and are possibly in the anti-sense orientation compared to the reference gene. This happens only when the alignment is so poor and when a sense or anti-sense alignment is nearly equivalent, and the anti-sense alignment was slightly better. In cases such as these, it is best to ignore the read entirely.

IMPORTANT: Use only reads from the *Merged Read Summary* to determine evidence for clonality.

Q9: Why is the analysis is taking a long time (over an hour per sample)?

A9: The log file provides a breakdown of how long each step is taking. Following the line ‘INFO: Decompressing files’, there is another line that states how many files and how much time the decompression used. This number will be in milliseconds. If decompression is taking several hours, check if the target files are on a fileserver. Typically, running any I/O intensive process like decompression over a network will be much slower. If possible, copy the files over to the local computer, or decompress the files directly before running this software. Make sure that the compressed and decompressed files do not both exist in the same folder. The software will not proceed if both *.fastq and *.fastq.gz files are in the target folder.

If the subsequent parsing and analysis steps are taking a long time, the *.fastq files may be very large. Examples of large *.fastq files would include when less than 24 indices are included on a single MiSeq run and as a result, the samples sequenced have a larger depth of coverage. A second example would be if the library is unbalanced and contains a higher proportion of a sample with one index versus the proportion of samples with other indexes. The predicted run-time stated in this IFU assumes a 24-index run with balanced libraries.

Q10: The Run Q30 reported by the MiSeq instrument fails the Illumina specification. If the LymphoTrack analysis is run and the controls look like they performed as expected, did the run really fail?

A10: The Run Q30 is calculated in real time during the base-calling process. At this point, no other data processing has taken place, such as de-multiplexing and adapter trimming. This means that any sequence from a poor-quality index or occurring outside of the adapter sequence is contributing to the Q30 score. Sequence found outside of the adapter sequence is typically low quality, and this scenario happens often with primer dimers. If a library has a sufficient level of primer dimers, this can cause a drop in Run Q30. As a secondary check, the LymphoTrack Reporter output displays the index Q30 score as calculated from all trimmed reads in that index. If an index Q30 score also fails the Illumina specification, consider that index to have failed.

Q11: What is the resolution for a failure when launching the LymphoTrack Dx Software – MiSeq by double-clicking the *startLymphoTrackDxMiSeq.jar* file?

A11: This issue can occur when there is a problem with the current installation of Java. The correct version of Java (Java 8 64-bit) must be installed on the system for the LymphoTrack Dx Software – MiSeq to run. Additionally, the *startLymphoTrackDxMiSeq.jar* file must be set to be opened with the correct version of Java.

The following troubleshooting steps can be taken to ensure that the correct version of Java is installed, and that the JAR is opened with the correct version of Java:

1. Verify the Java software is installed:
 - 1.1. Open the *Windows Menu* and type **Settings** in the search bar, then select the first icon that displays.
 - 1.2. In the search bar of the new window, type **Add or remove programs**, and select the first option that appears.
 - 1.3. Scroll through the list of installed programs.
 - If a program with a name starting with the letters *Java 8* and ending with the characters *(64-bit)* is in the list, the correct version of Java is installed.
 - If not, proceed to section 3 for instructions to install Java 8 64-bit.
2. Uninstall old versions of Java:
 - 2.1. Follow steps 1.1 and 1.2.
 - 2.2. Scroll through the list of installed programs. Identify any program that begins with the letters *Java* and also contains the letters *Java 8* and *64-bit*, click on the entry in the list, then click **Uninstall**.
3. Uninstall Java and Reinstall Java:

If the correct version of Java is installed, and the software still does not open, uninstall and re-install Java.

 - 3.1. Follow steps 1.1 and 1.2.
 - 3.2. Scroll through the list of installed programs and click on the entry with the name *Java 8* and ending with the characters *(64-bit)*, then click **Uninstall**.
 - 3.3. Navigate to the Java website to download the **Java 8 64-bit** software directly from Oracle at <http://www.java.com/>.
 - 3.4. Follow the instructions to install the software.
4. Point the JAR file to the correct Java executable:
 - 4.1. Right click on the *startLymphoTrackDxMiSeq.jar* file, then select **Properties**.
 - 4.2. In the dialog that appears, select the **Change** button next to the *opens with* text.
 - 4.3. In the next dialog that appears, select *Choose another app*.
 - 4.4. In the new window that opens, click on **More apps**, then select **Look for another app on this PC**.
 - 4.5. In the browsing window that appears, navigate to the folder location of the Java 8 64-bit installation.
 - This will likely be: **Local Disk (C:) > Program Files > Java > jre1.8.0_(version number) > bin**.
 - 4.6. Select the **javaw.exe** and click **Open**.
 - 4.7. Once installation is confirmed, repeat section 4.1 through 4.3 and select the checkbox indicating *Always open .jar files with this app*, then navigate to the same *javaw.exe* file from step 4.6. This will allow the LymphoTrack Dx Software to open correctly.

8. End User License Agreement

THIS END USER LICENSE AGREEMENT MUST BE ACCEPTED BY AN AUTHORIZED REPRESENTATIVE OF THE END USER OF THIS PRODUCT PRIOR TO USING THE LYMPHOTRACK® MISEQ SOFTWARE (the "Software"). BY USING THE LYMPHOTRACK® MISEQ SOFTWARE, YOU ACKNOWLEDGE THAT YOU ARE AN AUTHORIZED REPRESENTATIVE OF THE END USER WITH AUTHORITY TO ENTER INTO THIS AGREEMENT. PLEASE READ THIS AGREEMENT CAREFULLY. BY USING THE SOFTWARE, YOU ARE AGREEING TO BE BOUND BY THE TERMS OF THIS AGREEMENT. IF YOU DO NOT AGREE TO THE TERMS OF THIS AGREEMENT, PLEASE DISCONTINUE USAGE. IF YOU DO AGREE TO THE TERMS OF THIS AGREEMENT ON BEHALF OF YOURSELF AND/OR THE ENTITY YOU REPRESENT YOU MAY CONTINUE USAGE.

This End User License Agreement ("EULA") is made and entered into by and between INVIVOSCRIBE, INC., a California corporation ("Licensor") and you as the user of the Software (either you as an individual or a legal entity) ("Licensee") for the licensing and usage of the Licensor's Software. Licensee acknowledges and agrees that Licensee's right to use the Software in any manner shall be controlled by this EULA and that such use shall be strictly in accordance with the terms and conditions of this EULA.

1. **GRANT OF LICENSE.** Licensor grants to Licensee a non-exclusive, non-transferable, limited license, without the right to sublicense, to use the Software and any accompanying written materials and any documents or other content embodying information contained within the Software (collectively, the "Product") solely for Licensee's own internal use. Except as agreed in a document signed by Licensor, the software license granted in the preceding sentence is limited to use with Licensor's LymphoTrack® Assays purchased from Licensor or from an authorized distributor of Licensor. Licensor reserves all rights not expressly granted to Licensee or to Licensee. The limited license granted by this EULA and Licensee's payment of the license fee give Licensee the right to use the Product only in accordance with the terms of this EULA. This license is not a sale of the original software or any copy thereof.
2. **CONFIDENTIALITY.** Licensee agrees that the Product is owned by Licensor and is based on and includes the intellectual property of Licensor, including but not limited to one or more proprietary trade secrets, copyrights, patent applications, and/or granted patents ("Intellectual Property") of Licensor. Licensee's possession and use of the Product shall be strictly in accordance with this EULA, and receipt or possession does not convey any rights to divulge, reproduce, or allow others to use the Product outside the terms of this EULA without the specific written authorization of Licensor. Licensee agrees not to disclose, publish, translate, release, or distribute copies of the Product or any portion thereof to others outside the terms of this EULA. Licensee may not modify, adapt, translate, reverse engineer, decompile, disassemble, or create derivative works based on the Product or any portion thereof, including any documents or other content produced using the software.
3. **OWNERSHIP OF PRODUCT.** Title to, ownership of, and all rights and interests in the Product, including but not limited to the Software, all copies thereof, and "Intellectual Property" relating thereto, remains at all times vested in Licensor, and this EULA does not grant any ownership rights to Licensee.
4. **COPY RESTRICTIONS.** The Product is copyrighted. Unauthorized copying or modification of the Product, including software that has been modified, merged, or included with other software, is expressly forbidden. Licensee may be held legally responsible for any copyright infringement that is caused or incurred by Licensee's failure to abide by the terms of this EULA.
5. **TRANSFER RESTRICTIONS.** This Product is licensed to Licensee, and may not be transferred to anyone without the prior written consent of Licensor. In no event may Licensee transfer, assign, rent, lease, sell, or otherwise dispose of the Product, or any portion thereof, on a temporary basis, except as expressly provided herein.
6. **TERMINATION.** This EULA will be terminated automatically without notice if Licensee fails to comply with any of its provisions.
7. **LIMITED WARRANTY.** THE PRODUCT, INCLUDING SOFTWARE AND ACCOMPANYING WRITTEN MATERIALS, INCLUDING INSTRUCTIONS FOR USE, ARE PROVIDED "AS IS," WITHOUT WARRANTY OF ANY KIND AND LICENSOR DISCLAIMS ALL OTHER WARRANTIES, WHETHER EXPRESS, IMPLIED, OR STATUTORY, INCLUDING ANY IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE, TITLE, NON-INFRINGEMENT AND ANY WARRANTIES ARISING FROM COURSE OF DEALING OR COURSE OF PERFORMANCE) REGARDING OR RELATING TO THE SOFTWARE, THE DOCUMENTATION, OR ANY MATERIALS FURNISHED OR PROVIDED TO YOU UNDER THIS EULA. LICENSOR DOES NOT WARRANT THAT THE SOFTWARE WILL OPERATE UNINTERRUPTED OR THAT IT WILL BE FREE FROM DEFECTS OR THAT THE SOFTWARE WILL MEET (OR IS DESIGNED TO MEET) YOUR BUSINESS REQUIREMENTS.
8. **LIMITED LIABILITY.** TO THE MAXIMUM EXTENT PERMITTED BY LAW, IN NO EVENT WILL LICENSOR BE LIABLE FOR ANY LOST PROFITS OR BUSINESS OPPORTUNITIES, LOSS OF USE, LOSS OF REVENUE, LOSS OF GOODWILL, BUSINESS INTERRUPTION, LOSS OF DATA, OR ANY OTHER INDIRECT, SPECIAL, INCIDENTAL, OR CONSEQUENTIAL DAMAGES UNDER ANY THEORY OF LIABILITY, WHETHER BASED IN CONTRACT, TORT, NEGLIGENCE, PRODUCT LIABILITY, OR OTHERWISE. BECAUSE SOME JURISDICTIONS DO NOT ALLOW THE EXCLUSION OR LIMITATION OF LIABILITY FOR CONSEQUENTIAL OR INCIDENTAL DAMAGES, THE PRECEDING LIMITATION MAY NOT APPLY TO LICENSEE. LICENSOR'S LIABILITY UNDER THIS EULA WILL NOT, IN ANY EVENT, REGARDLESS OF WHETHER THE CLAIM IS BASED IN CONTRACT, TORT, STRICT LIABILITY, OR OTHERWISE, EXCEED THE LICENSE FEES LICENSEE PAID FOR THE SOFTWARE, IF ANY. THE FOREGOING LIMITATIONS SHALL APPLY REGARDLESS OF WHETHER LICENSOR HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES AND REGARDLESS OF WHETHER ANY REMEDY FAILS OF ITS ESSENTIAL PURPOSE.
9. **EXCLUSIVE REMEDY.** Licensee's exclusive remedy and Licensor's entire liability for breach by Licensor of this EULA shall be limited, at Licensor's sole and exclusive discretion, to: (i) replacement of any defective Software or related Documentation; or (ii) refund of the consideration paid to Licensor for the Software.
10. **GOVERNING LAW.** This EULA shall be governed by the laws of the State of California, irrespective of its conflicts of laws principles, and Licensee consents to jurisdiction by the state and federal courts sitting in the State of California.
11. **COMPLETE AGREEMENT.** This EULA is the entire agreement between Licensor and Licensee with respect to the specific terms set forth herein concerning license and warranties of the Product and any other included term or obligation. This EULA replaces all prior understandings and agreements, whether written or oral. This EULA may not be modified unless Licensor and Licensee both agree to such modification in writing.
12. **SEVERABILITY.** If for any reason a court of competent jurisdiction finds any provision or part of any provision of this EULA unenforceable, that part or provision shall be enforced to the maximum extent permitted by law so as to affect the intent of the parties, and the remainder of the EULA shall continue in full force and effect.

©2024 Invivoscribe, Inc. All rights reserved. The trademarks mentioned herein are the property of Invivoscribe, Inc. and/or its affiliates, or (as to the trademarks of others used herein) their respective owners.