LymphoTrack® Dx Assays

Instructions d'utilisation

CE UKCA IVD

LymphoTrack® Dx Software – MiSeq™ Version 2.4.6

Cette suite logicielle est fournie avec chaque test LymphoTrack Dx Assay pour l'instrument Illumina MiSeq afin d'analyser des fichiers FASTQ bruts. Le logiciel se compose de trois parties distinctes : l'application d'analyse de données LymphoTrack, un plugin LymphoTrack Reporter qui est exécuté automatiquement au terme de l'analyse, ainsi qu'une méthode de visualisation LymphoTrack facultative qui consiste en quatre feuilles de calcul Microsoft Excel® XLSM. Le logiciel fonctionne sur les systèmes Microsoft Windows® pris en charge.

Cette suite logicielle est conçue pour l'analyse de la clonalité d'un ou de plusieurs gènes cibles (*IGHV* Leader [ou « Leader »], *IGH* FR1, *IGH* FR2, *IGH* FR3, *IGK, TRB* ou *TRG*) et indique également le statut d'hypermutation somatique des échantillons analysés avec les tests Leader ou *IGH* FR1.

IVD Destiné au diagnostic *in vitro*.

Compatibilité du produit

Le logiciel LymphoTrack Dx Software – MiSeq (REF 95000009) a été développé spécifiquement et **uniquement** pour les tests d'Invivoscribe suivants :

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

invivoscribe

Table des matières

1. UTILISATION PREVUE 3
2. GLOSSAIRE ET ABREVIATIONS 3
3. SYSTEME MINIMAL REQUIS 4
4. MISES EN GARDE ET PRECAUTIONS 4
5. INSTALLATION DU LOGICIEL 4
6. EXECUTION DU LOGICIEL D'ANALYSE DE DONNEES 5
7. VISUALISATION DES DONNEES A L'AIDE DE LYMPHOTRACK REPORTER (AUTOMATISEE, RECOMMANDEE) 9
8. VISUALISATION DES DONNEES (MANUELLE) : EXECUTION DES FEUILLES DE CALCUL DE VISUALISATION DES DONNEES EXCEL - DECONSEILLEE 10
9. SUPPORT TECHNIQUE ET SERVICE CLIENT 13
10. SYMBOLES 13
11. INFORMATIONS LEGALES 14
12. FOIRE AUX QUESTIONS CONCERNANT LE LOGICIEL 14

1. Utilisation prévue

La suite logicielle LymphoTrack Dx Software – MiSeq est prévue pour être utilisée à des fins d'analyse des données bio-informatiques et de visualisation des fichiers FASTQ bruts de l'instrument Illumina MiSeq lors de l'analyse conjointe de tests LymphoTrack Dx Assays.

2. Glossaire et abréviations

2.1 Glossaire

Tableau 1. Termes du glossaire

Terme	Définition
Clonal	L'agrégat de cellules ou d'organismes génétiquement identiques produits à partir d'une seule cellule progénitrice.
Invalide	Résultat d'un échantillon ou d'un contrôle qui ne répond pas aux critères de validité.
Versionnage sémantique	Un schéma de version de logiciel composé de 3 chiffres (Major.Minor.Patch), aligné sur le facteur de risque de la mise à jour.
Logiciel	Logiciel LymphoTrack Dx - MiSeq
Système	L'ensemble des logiciels, du matériel et de l'appareillage (le cas échéant) qui composent le dispositif médical.

2.2 Abréviations

Tableau 2. Abréviations définies

Abréviation	Définitions
CDR3	Région déterminant la complémentarité, formée par la jonction des segments V-(D)-J et établissant la diversité de la réponse immunitaire cellulaire.
CSV	Format de fichier des valeurs séparées par des virgules
FASTQ	Format textuel permettant de stocker à la fois une séquence biologique (généralement une séquence de nucléotides) et les scores de qualité correspondants.
FFPE	Fixés à la formaline et inclus dans la paraffine
IFU	Instructions d'utilisation
***IGH* FR1**	Chaîne lourde d'immunoglobulines Cadre 1
***IGH* FR2**	Chaîne lourde d'immunoglobulines Cadre 2
***IGH* FR3**	Chaîne lourde d'immunoglobulines Cadre 3
IGHV	Du gène des régions variables des chaînes lourdes de l'immunoglobuline
IGK	Chaîne légère de l'immunoglobuline kappa
PCR	Réaction en chaîne par polymérase
PDF	Format de document portable
SHM	Hypermutation somatique
TRB	Récepteur bêta des lymphocytes T
TRG	Récepteur gamma des lymphocytes T
TSV	Format du fichier de valeurs séparées par des tabulations
USB	Bus universel en série ; permet l'échange de données et la fourniture d'énergie entre de nombreux types d'appareils électroniques.

3. Système minimal requis

- **Processeur :** Intel Core 2 Duo ou processeur plus récent recommandé.
- **Disque dur :** au moins 80 Go d'espace disque disponible sont nécessaires ; 250 Go sont recommandés.
- **Mémoire RAM :** 4 Go nécessaires ; 8 Go ou plus sont recommandés.
- **Système d'exploitation :** Windows 10 Pro ou Windows 11 Pro (64 bits) est nécessaire.
- Un lecteur de fichiers PDF, tels qu'Adobe Acrobat Reader, pour visualiser les rapports de données générés par LymphoTrack Reporter.
- Microsoft Office Excel 2007, 2010 ou 2013 avec exécution de macros autorisée est nécessaire pour la visualisation des données.
- Un lecteur de CD-ROM.
- Le logiciel LymphoTrack Dx Software – MiSeq v2.4.6.IVD a été validé pour une utilisation avec Windows 10 Pro et Windows 11 Pro configuré avec les paramètres de langue Anglais (États-Unis) et de taille d'affichage par défaut, ainsi que Microsoft Excel 2007/2010/2013 pour obtenir une visualisation différente des données. L'utilisation d'autres versions de Microsoft Excel/Windows et/ou paramètres de langue/taille peut ne pas être compatible.

4. Mises en garde et précautions

- **Taille d'écran par défaut.** La résolution d'écran optimale pour LymphoTrack Dx Software – MiSeq est 100 %. L'augmentation de la taille d'écran par défaut du moniteur à plus de 100 % peut affecter l'affichage du texte des champs dans l'interface utilisateur.
- **Police système**. L'interface utilisateur est conçue pour utiliser les paramètres de police par défaut du système sur un ordinateur Windows utilisant les paramètres régionaux Anglais (É.-U.). Les changements apportés aux paramètres de police du système ou l'utilisation de paramètres régionaux de langue autres que l'anglais (É.-U.) peuvent avoir une incidence sur l'affichage du texte des champs dans l'interface utilisateur.
- **Fichiers compatibles.** LymphoTrack Dx Software – MiSeq est compatible avec les fichiers FASTQ.GZ générés par le logiciel MiSeq Reporter. Le logiciel LymphoTrack Dx Software – MiSeq n'est pas compatible avec les fichiers FASTQ.GZ générés par le logiciel *Bcl2fastq*.
- **Élimination des adaptateurs.** L'option Adapter trimming (Élimination des adaptateurs) doit être sélectionnée ☑ lors de la création des feuilles d'échantillons MiSeq afin que les fichiers FASTQ.GZ générés ne contiennent pas d'adaptateurs.
- **Paramètres de veille ou de veille prolongée.** Penser à désactiver le mode veille et/ou veille ou veille prolongée avant d'exécuter le logiciel d'analyse de données LymphoTrack Dx Data Analysis Application. Si l'ordinateur entre en mode veille ou veille prolongée, l'exécution du logiciel risque d'être interrompue.
- **Caractères utilisés pour les noms de chemin et de fichier.** 1) Éviter d'inclure des espaces dans le nom du chemin des fichiers à analyser (les noms de chemin comprennent les noms de dossier et les noms de fichier) et ne pas utiliser plus d'un espace consécutif. 2) Il est important que les noms de fichier contiennent uniquement des caractères alphanumériques et tirets suivants (A-Z, a-z, 0-9, - ([tiret[). Si le logiciel rencontre un caractère non compris dans cet ensemble, son exécution peut échouer. Veillez à suivre cette nomenclature d'affectation de nom pour les échantillons lors de la configuration de la feuille d'échantillons pour le MiSeq. 3) Le nom du chemin ne doit pas dépasser 255 caractères. Si, en raison des noms des échantillons, les noms de fichiers générés avoisinent cette limite, ces derniers seront difficiles à copier vers d'autres emplacements.
- **Les échantillons en double doivent comporter un nom d'échantillon unique.** Attribuer un nom et un identifiant uniques à chaque échantillon lorsque vous les nommez. Si les échantillons sont analysés en double, ils peuvent être nommés « échantillon1a » et « échantillon1b ». Si des noms uniques ne sont pas attribués aux échantillons analysés ensemble dans la même Flow Cell, un seul échantillon est analysé par le logiciel au cours du processus.

5. Installation du logiciel

5.1 Double-cliquez sur l'archive ***LymphoTrackDxMiSeq-2.4.6.IVD.msi*** situé sur le CD du logiciel et suivre les instructions à l'écran.

- Installez uniquement le logiciel sur un disque local et non sur un lecteur réseau.
- L'exécution du logiciel au moyen d'une connexion réseau peut ne pas fonctionner correctement.

6. Exécution du logiciel d'analyse de données

6.1 Identifiez les fichiers FASTQ.GZ à lecture appariée, avec adaptateurs éliminés, à analyser. Ces derniers se trouvent dans le dossier d'analyse MiSeq et sont des fichiers FASTQ.GZ qui seront utilisés comme source pour l'application d'analyse de données **LymphoTrackDxMiSeq**.

- Les fichiers *Undetermined_*.fastq.gz* ne sont pas nécessaires pour l'analyse.

6.2 Copiez les fichiers FASTQ.GZ identifiés ci-dessus dans un dossier vide de l'ordinateur. Téléchargez ces fichiers sur le même ordinateur local que celui sur lequel le logiciel LymphoTrack Dx Software est installé.

- Le disque contenant ce dossier doit comporter au moins 80 Go d'espace libre.
- Ce dossier sera nommé **<input_folder>** (**<dossier_source>**) comme expliqué ci-dessous pour le logiciel.

6.3 Double-cliquez sur le fichier de l'application ***LymphoTrackDxMiSeq-2.4.6.IVD*** situé dans le dossier **Invivoscribe** créé lors de l'installation.

- L'emplacement par défaut est *Disque local (C:) > Invivoscribe*

6.4 Veuillez lire le contrat de licence. Pour accepter les conditions du contrat, cliquez sur le bouton **Accept** (**Accepter**) pour continuer.

6.5 Cochez le mode d'analyse souhaité dans les options de la case à cocher *Choose analysis mode(s) (Choisir le(s) mode(s) d'analyse)*.

- Le mode d'analyse doit correspondre au type de données dans l'analyse (Leader, *IGH* FR1, *IGH* FR2, *IGH* FR3, *IGK*, *TRB* et/ou *TRG*).

6.6 Sélectionnez le séparateur décimal souhaité parmi les options *Choose decimal mark (Choisir un séparateur décimal)*.

- Cette option a des conséquences sur l'utilisation de la macro de visualisation des données dans Microsoft Excel par la suite et nécessite d'être renseignée.
- Les choix possibles sont un séparateur décimal sous forme de point (.) ou de virgule (,).

IMPORTANT ! **Séparateur décimal.** Pour déterminer quelle option choisir, ouvrez une nouvelle feuille de calcul Excel et saisissez « **=10000 »** dans la cellule A1. Effectuez un clic droit sur cette cellule et sélectionnez **Format Cells** (Format de cellule).

- ☐ Dans l'onglet *Number* (*Nombre*), passez en revue la liste *Category:* (*Catégorie* :) et sélectionnez **Number** (**Nombre**). Sur la droite, la valeur *Decimal places* (*Nombre de décimales*) doit être définie sur « **2** ».
- ☐ En dessous se trouve la case à cocher *Use 1000 Separator* (*Utiliser le séparateur de milliers*). Cochez cette case et observez l'exemple de nombre au-dessus. Le nombre original que vous avez saisi doit avoir le format correspondant aux paramètres d'Excel par défaut pour un nombre comprenant un séparateur décimal.
- ☐ Si ce nombre apparaît sous la forme « *10,000.00* », sélectionnez l'option **decimal point** (**point comme séparateur décimal**). Si ce nombre apparaît sous la forme « *10.000,00* », sélectionnez l'option **decimal comma** (**virgule comme séparateur décimal**).

6.7 Cliquez sur le bouton **Browse** (**Parcourir**) pour choisir un dossier source. Ce dossier est le dossier **<input_folder>** (**<dossier_source>**) mentionné ci-dessus et contient les fichiers à analyser. Lorsque le dossier source est sélectionné, il semblera vide.

IMPORTANT ! **Le <input_folder> sélectionné ne doit pas contenir de fichiers échantillon portant le même nom que le dossier source**. Si le programme a été exécuté avec un échantillon partageant le même nom de fichier que le dossier fourni en entrée de la fonctionnalité d'analyse, le programme se bloquera et le temps d'exécution de l'analyse augmentera considérablement (c'est-à-dire qu'il faudra 30 minutes pour traiter un petit échantillon qui prend quelques secondes à traiter normalement).

6.8 Une fois le **<input_folder>** et le mode d'analyse sélectionnés, cliquez sur le bouton **Launch Program** (**Démarrer le programme**).

IMPORTANT ! **Le <input_folder> sélectionné doit contenir uniquement des fichiers FASTQ ou FASTQ.GZ.** En l'absence de ces types de fichier (seuls ou combinés), le programme affiche un message d'erreur et demande de sélectionner un autre dossier. Si les fichiers FASTQ ont été générés avec *Bcl2fastq* (les noms de fichier comportent la séquence du code-barre à la place de la nomenclature « S1, S2, S3 » etc.), le programme affiche un message d'erreur et demande d'analyser de nouveau les données à l'aide du logiciel MiSeq Reporter.

6.9 **État de l'analyse**. Après avoir cliqué sur le bouton **Launch Program** (**Démarrer le programme**), la fenêtre principale du programme affiche l'état du programme. Lorsque le programme a fini d'analyser les fichiers de données sources, la fenêtre principale du programme affiche le message *Analysis is complete…* (*L'analyse est terminée*...) et indique l'emplacement du ou des dossiers de sortie en invitant à ouvrir la feuille de calcul Excel de visualisation des données correspondante.

- Une nouvelle analyse peut être effectuée à partir de la fenêtre principale du programme dès que la précédente est terminée.

IMPORTANT ! **Durée de l'analyse.** La taille et le nombre d'échantillons à analyser, ainsi que le mode d'analyse et le matériel informatique déterminent la durée d'exécution du programme. Une distribution type pour une analyse de 24 échantillons est de 200 000 à 1 million de lectures par échantillon. Ces chiffres peuvent varier avec les versions ultérieures de la plate-forme MiSeq.

L'analyse d'un échantillon type unique des gènes *IGHV* Leader, *IGH* FR1, *IGH* FR2, *IGH* FR3, *IGK, TRB* ou *TRG* nécessite environ 1 à 2 minutes. L'analyse de 24 échantillons types des gènes Leader seul, *IGH* FR1 seul, *IGH* FR2 seul, *IGH* FR3 seul, *IGK* seul, *TRB* seul ou *TRG* seul nécessite environ 25 à 50 minutes. La durée d'analyse de 24 échantillons types combinés des gènes Leader, *IGH* FR1, *IGH* FR2, *IGH* FR3, *IGK*, *TRB* et/ou *TRG* est fonction du nombre de cibles à analyser.

6.10 **Dossier(s) de sortie.** En fonction du mode d'analyse choisi, un ou plusieurs dossiers de sortie seront créés dans le **<input_folder>** contenant les fichiers sources FASTQ ou FASTQ.GZ.

- Le Tableau 3 ci-dessous présente la relation entre le mode d'analyse et le nom du ou des dossiers de sortie.

Tableau 3. Noms du dossier de sortie en fonction du mode d'analyse

Mode d'analyse	Noms du ou des dossiers de sortie
Leader	*LEADER_output*
IGH FR1	*IGH_FR1_output*
IGH FR2	*IGH_FR2_output*
IGH FR3	*IGH_FR3_output*
TRB	*TRB_output*
TRG	*TRG_output*
IGK	*IGK_output*
Toute analyse combinée de : Leader, *IGH* FR1, *IGH* FR2, *IGH* FR3, *IGK*, *TRB* et *TRG*	Deux à sept dossiers de sortie énumérés ci-dessus

6.11 Lorsque le programme a terminé l'analyse avec succès, le dossier de sortie contient plusieurs fichiers TSV et un fichier PDF pour chaque échantillon traité.

- Remarque : dans les tableaux ci-dessous, « * » représente le nom de l'échantillon tel qu'il est saisi dans la feuille d'échantillons, suivi des informations spécifiques de la plateforme (*p. ex*, nom d'échantillon_S1_L001_001_ combiné).
- Pour l'analyse des gènes *IGH* FR1, *IGH* FR2, *IGH* FR3, *TRB* et *IGK*, veuillez consulter le Tableau 4 présentant les descriptions des neuf fichiers fournis pour chaque échantillon.

Tableau 4. Description des fichiers de sortie *IGH*, *TRB* et *IGK*.

Analyse des gènes Leader, *IGH* FR1, *IGH* FR2, *IGH* FR3, *TRB* et *IGK*	
**.fastq_read_summary.tsv*	Contient une liste des 200 séquences uniques les plus importantes, triées par nombre (nombre de lectures) et précisant : ▪ À quel événement de recombinaison V-J elle correspond le mieux ▪ La longueur ▪ Le % de lectures totales réel ▪ Le % de lectures totales cumulé ▪ Dans quelle mesure la lecture couvre le gène V identifié ciblé par les amorces (%) (*IGH/IGK* et *TRB*) ▪ La séquence putative CDR3 Leader et *IGH* FR1 précisent également : ▪ Le taux de mutations du gène V partiel (%) Leader, *IGH* FR1, *IGH* FR2, *IGH* FR3 précisent également : ▪ Si le réarrangement doit se situer à l'intérieur du cadre (Y/N/n/a [O/N/s.o.]). Si l'utilisation du segment de gène V ou J ne peut pas être prédite en raison de la qualité de l'alignement, le logiciel ne sera pas en mesure de prédire si le gène se situe à l'intérieur du cadre et le résultat sera n/a. ▪ Si le réarrangement ne devrait pas contenir un codon stop (Y/N [O/N]). ▪ Si la position à l'intérieur du cadre ne peut pas être prédite, il n'est pas possible de prédire avec exactitude si un codon stop sera présent.
**.fastq_read_summary_merged_top10_ searchtop500.tsv*	Même format que le fichier *read_summary*, mais il contient uniquement les 10 premières lectures du fichier *read_summary* fusionnées avec les 500 premières lectures. Une lecture est fusionnée avec une autre si seulement 1 ou 2 nucléotides (nts) les différencient l'une de l'autre ; la lecture comportant le nombre le plus élevé conserve l'identité de sa séquence et la lecture comportant le nombre le moins élevé est ajoutée à la plus importante.
**.fastq_unique_reads.tsv*	Fichier FASTA comportant des séquences uniques, triées en fonction de leur nombre.
**.fastq_CDR3_unique_reads.tsv*	Fichier FASTA de séquences putatives CDR3 figurant dans le fichier FASTQ_UNIQUE_READS.
**.fastq_VJ_sequence_frequency_family.tsv*	Chaque famille regroupe toutes les recombinaisons possibles que l'on trouve dans une région V ou J particulière, comme la famille V_1 ou la famille J_4. On peut trouver n'importe quel nombre de recombinaisons au sein de la famille V_1, tels que $V_{1\text{-}69}$ et $V_{1\text{-}8}$. Ce fichier contient une liste de toutes les recombinaisons possibles de la famille V – famille J, chacune comportant une liste de lectures triées (du % le plus élevé au plus faible). Ce fichier est parfaitement adapté aux graphiques « empilés » dans Excel.
**.fastq_read_summary_family.tsv*	Contient une liste triée de séquences uniques, par nombre et précisant : ▪ À quel événement de recombinaison de la famille V-J elle correspond le mieux ▪ La longueur ▪ Le % de lectures totales réel ▪ Le % de lectures totales cumulé ▪ Dans quelle mesure la lecture couvre le gène V identifié ciblé par les amorces (%) (*IGH/IGK*) ▪ La séquence putative CDR3 Leader et *IGH* FR1 précisent également : ▪ Le taux de mutations du gène V partiel (%) Leader, *IGH* FR1, *IGH* FR2, *IGH* FR3 précisent également : ▪ Si le réarrangement doit se situer à l'intérieur du cadre (Y/N/n/a [O/N/s.o.]) ▪ Si le réarrangement ne devrait pas contenir un codon stop (Y/N [O/N])
**.fastq_VJ_usage.tsv*	Matrice de numération et matrice de % des différents événements de recombinaison V-J.
**.fastq_VJ_usage_family.tsv*	Matrice de numération et matrice de % des différents événements de recombinaison famille V – famille J.

Tableau 4. Description des fichiers de sortie *IGH, TRB* et *IGK*.

Analyse des gènes Leader, *IGH* FR1, *IGH* FR2, *IGH* FR3, *TRB* et *IGK*	
**<horodatage>.pdf*	Le rapport PDF généré par LymphoTrack Reporter comporte des informations semblables à celles des fichiers *merged read summary (Résumé de lecture fusionné)* et *read summary (Résumé de lecture)* ainsi que des graphiques *V – J usage (Utilisation de VJ)* et *V – J sequence frequency (Fréquence de la séquence VJ)*. Voir la rubrique 2.13 *Visualisation des données à l'aide de LymphoTrack Reporter (Automatisée, Recommandée)* pour plus d'informations.

Remarque : *l'analyse d'IGK* inclut également des informations sur d'autres éléments du locus *IGK* impliqués dans les recombinaisons (IGKINTR et IGKDEL).
Remarque : *l'analyse de TRB* inclut également des informations sur d'autres éléments du locus *TRB* impliqués dans les recombinaisons de la région D – J.
Remarque : *la séquence putative CDR3* est déterminée par la correspondance d'une définition exacte d'une séquence CDR3. Toute mutation des points d'ancrage ou perturbation du cadre de lecture entraînera une valeur *not found (non trouvé)*.

- Pour l'analyse de *TRG*, veuillez consulter le Tableau 5 présentant les descriptions des sept fichiers fournis pour chaque échantillon.

Tableau 5. Description des fichiers de sortie *TRG*.

Analyse de *TRG*	
**.fastq_read_summary.tsv*	Une liste des séquences uniques triées par nombre et précisant : ▪ À quel événement de recombinaison V-J elle correspond le mieux ▪ La longueur ▪ Le % de lectures totales réel ▪ Le % de lectures totales cumulé ▪ La séquence putative CDR3
**.fastq_read_summary_merged_top10_searchtop500.tsv*	Même format que le *fichier read_summary*, mais il contient uniquement les 10 premières lectures du fichier *read_summary* fusionnées avec les 500 premières lectures. Une lecture est fusionnée avec une autre si seulement 1 ou 2 nts les différencient l'une de l'autre ; la lecture comportant le nombre le plus élevé conserve l'identité de sa séquence et la lecture comportant le nombre le moins élevé est ajoutée à la plus importante.
**.fastq_unique_reads.tsv*	Fichier FASTA comportant des séquences uniques, triées en fonction de leur nombre.
**.fastq_CDR3_unique_reads.tsv*	Fichier FASTA de séquences putatives CDR3 figurant dans le fichier FASTQ_UNIQUE_READS.
**.fastq_VJ_sequence_frequency.tsv*	Liste de toutes les recombinaisons V-J possibles, chacune comportant une liste de lectures triées (du % le plus élevé au plus faible). Le premier nombre représente le % de lectures totales que la première lecture unique a comptabilisé sur toutes les lectures de séquences. Ce fichier est parfaitement adapté aux graphiques « empilés » dans Excel.
**.fastq_VJ_usage.tsv*	Matrice de numération et matrice de % des différents événements de recombinaison V-J.
**<horodatage>.pdf*	Le rapport PDF généré par LymphoTrack Reporter comporte des informations semblables à celles des fichiers *merged read summary (Résumé de lecture fusionné)* et *read summary (Résumé de lecture)* ainsi que des graphiques *VJ usage (Utilisation de V–J)* et *VJ sequence frequency (Fréquence de la séquence V–J)*. Voir la rubrique 2.13 *Visualisation des données à l'aide de LymphoTrack Reporter (Automatisée, Recommandée)* pour plus d'informations.

Remarque : *la « séquence putative CDR3 »* est déterminée par la correspondance d'une définition exacte d'une séquence CDR3. Toute mutation des points d'ancrage ou perturbation du cadre de lecture entraînera une valeur « not found » (non trouvé).

- Le(s) dossier(s) de sortie contiennent des rapports PDF et des fichiers de données de sortie TSV.
- Les fichiers TSV peuvent être utilisés comme source de données pour la ou les feuilles de calcul de visualisation des données.

7. Visualisation des données à l'aide de LymphoTrack Reporter (Automatisée, Recommandée)

7.1 Un rapport PDF est automatiquement généré pour chaque échantillon à l'issue de l'analyse.

- Le rapport est nommé en fonction du nom de l'échantillon présent dans la feuille d'échantillons, suivi des informations spécifiques de la plateforme, et se termine par des informations d'horodatage (JJMMAAhhmmss).

IMPORTANT ! **Lors de l'enregistrement du rapport PDF:** en fonction du programme utilisé pour ouvrir le PDF, il est possible de modifier une ou plusieurs colonnes. Ne pas modifier une quelconque partie du PDF. Si l'utilisateur suspecte qu'une modification a été effectuée, ne pas enregistrer le PDF . Si l'utilisateur suspecte qu'une modification a été enregistrée, relancer l'analyse afin de générer un nouveau rapport PDF.

7.2 La première page du rapport PDF comprend une valeur ***Index Q30***, qui correspond au pourcentage des bases ayant un score de qualité de 30 ou plus, mais provenant uniquement de lectures affectées à un index après élimination. Cette valeur peut être utilisée pour déterminer si l'index répond à la spécification standard Illumina.

- Dans certains cas, le score de l'analyse Q30 fourni par l'outil de séquençage peut sous-estimer la qualité d'une analyse du fait de la prise en compte de séquences non déterminées et de régions non éliminées en vue de l'analyse Q30.
- Les régions non déterminées et non éliminées sont généralement de moins bonne qualité et ne sont pas pertinentes pour l'analyse en aval par LymphoTrack.

7.3 La première page du rapport PDF comprend un tableau des résultats du ***merged read summary*** ***(Résumé de lecture fusionné)***, suivi de graphiques représentant les résultats de ***V – J Sequence Frequency*** et de ***V – J Usage*****. Les pages restantes** contiennent un tableau des résultats du ***read summary*** ***(Résumé de lecture)*** non fusionnés.

- Ces résultats sont détaillés dans le Tableau 6:

Tableau 6. Description du contenu du rapport PDF.

Rapport PDF de LymphoTrack Reporter	

Read Summary
(Résumé de lecture)
	Contient une liste des 200 séquences uniques les plus importantes, triées par nombre et précisant :
	• À quel événement de recombinaison V – J chaque séquence correspond le mieux
	  ○ La séquence
	  ○ La longueur de la séquence
	  ○ Le % de lectures totales réel
	• Le % de lectures totales cumulé sous forme de liste pour les 200 lectures les plus importantes
	  ○ Dans quelle mesure la lecture couvre le gène V identifié ciblé par les amorces (%) (*IGH/IGK*)
	  ○ La séquence putative CDR3
	
	Leader et *IGH* FR1 précisent également :
	• Le taux de mutations du gène V partiel (%)
	
	Leader, *IGH* FR1, *IGH* FR2, *IGH* FR3 précisent également :
	• Si le réarrangement doit se situer à l'intérieur du cadre (Y/N/n/a [O/N/s.o.]). Si l'utilisation du segment de gène V ou J ne peut pas être prédite en raison de la qualité de l'alignement, le logiciel ne sera pas en mesure de prédire si le gène se situe à l'intérieur du cadre et le résultat sera n/a.
	• Si le réarrangement ne doit pas contenir un codon stop (Y/N [O/N]). Si la position à l'intérieur du cadre ne peut pas être prédite, il n'est pas possible de prédire avec exactitude si un codon stop sera présent.

Merged Read Summary *(Résumé de lecture fusionné)*	Même format que la feuille de travail Read Summary (Résumé de lecture), mais uniquement pour les 10 premières lectures après fusion avec les 500 premières si les lectures ne se différencient l'une de l'autre que par 1 ou 2 nts. En outre, le score Q30 d'index est affiché sous Total Read Count (Nombre total de lectures).

Tableau 6. Description du contenu du rapport PDF.

Rapport PDF de LymphoTrack Reporter	
V – J Sequence Frequency Graph *(Graphique de la fréquence de la séquence V – J)*	Histogramme représentant les 200 fréquences de séquence les plus importantes pour chaque famille V – famille J (pour *IGH*/*IGK* et *TRB*) ou recombinaison de gène V – gène J (pour *TRG*).
V – J Usage Raw Graph *(Graphique brut d'utilisation de V – J)*	Graphique en barres 3D illustrant l'utilisation de la famille V – famille J des 200 séquences les plus importantes (pour *IGH*/*IGK* et *TRB*) ou l'utilisation du gène V – gène J pour toutes les séquences (pour *TRG*).

Remarque : *la séquence putative CDR3* est déterminée par la correspondance d'une définition exacte d'une séquence CDR3. Toute mutation des points d'ancrage ou perturbation du cadre de lecture entraînera une valeur « not found » (non trouvé).
Remarque : *l'analyse d'IGK* inclut également des informations sur d'autres éléments du locus *IGK* impliqués dans les recombinaisons (IGKINTR et IGKDEL).
Remarque : *l'analyse de TRB* inclut également des informations sur d'autres éléments du locus *TRB* impliqués dans les recombinaisons de la région D – J.

8. Visualisation des données (manuelle) : exécution des feuilles de calcul de visualisation des données Excel - déconseillée

8.1 Vérifiez les paramètres de sécurité dans Microsoft Excel afin de vous assurer que l'exécution des macros est bien activée :

8.1.1. Excel 2007 : cliquez sur le bouton **Office** et sélectionnez **Excel Options (Options Excel)**.
8.1.2. Excel 2010/2013 : cliquez sur le menu **File** (**Fichier**) et sélectionnez **Options**.
8.1.3. Dans la fenêtre *Options*, cliquez sur **Trust Center** (**Centre de gestion de la confidentialité**), puis sur le bouton **Trust Center Settings...** (**Paramètres du Centre de gestion de la confidentialité…**).
8.1.4. Cliquez sur **Macro Settings** (**Paramètres des macros**) et définissez le niveau de sécurité des macros approprié. L'option *Disable all macros with notification* (*Désactiver toutes les macros avec notification*) est le paramètre recommandé.
8.1.5. Cliquez sur **OK**, puis de nouveau sur **OK**.
8.1.6. Fermez Excel.

IMPORTANT! Étant donné que les fichiers de visualisation peuvent être enregistrés sous forme de classeur (ce qui supprimerait la capacité de la macro à fonctionner), copiez les fichiers de visualisation dans un chemin de fichier différent avant de commencer les analyses.

8.2 Double-cliquez sur la feuille de calcul de visualisation des données appropriée pour l'ouvrir.

- Les fichiers de visualisation sont situés dans le sous-dossier « visualisations » situé dans le dossier d'installation du logiciel. Les fichiers des feuilles de calcul de visualisation suivantes sont disponibles :
  - ☐ *LymphoTrackDxIGHVisualization.xlsm* pour Leader, *IGH* FR1, *IGH* FR2, *IGH* FR3
  - ☐ *LymphoTrackDxIGKVisualization.xlsm* pour *IGK*
  - ☐ *LymphoTrackDxTRBVisualization.xlsm* pour *TRB*
  - ☐ *LymphoTrackDxTRGVisualization.xlsm* pour *TRG*

8.3 À l'ouverture du fichier, le message suivant peut s'afficher : Security Warning - Macros have been disabled (Avertissement de sécurité - Les macros ont été désactivées).

8.3.1. Excel 2007 : cliquez sur le bouton **Options...** à droite de l'avertissement et sélectionnez **Enable this content** (Activer le contenu), puis cliquez sur **OK**.
8.3.2. Excel 2010/2013 : cliquez sur le bouton **Enable Content** (Activer le contenu).

- La macro de visualisation des données démarre automatiquement.

8.4 Le contrat de licence s'affiche. Il est nécessaire d'accepter le contrat de licence afin d'utiliser le logiciel.

- Après avoir accepté le contrat de licence, une boîte de dialogue s'ouvre avec le message suivant : *Please select the Read Summary file for the sample you want. (Veuillez sélectionner le fichier de résumé de lecture pour l'échantillon souhaité)*.
- La macro peut uniquement effectuer une visualisation des données sur un seul échantillon à la fois. Par conséquent, cette procédure devra être répétée pour chaque échantillon à visualiser.

8.5 Parcourez le fichier de sortie créé par l'application d'analyse des données contenant les données à visualiser et sélectionnez le fichier ***FASTQ_READ_SUMMARY*** pour l'échantillon à étudier.

8.6 Cliquez sur le bouton **Open** (**Ouvrir**).

- La macro charge les fichiers de données pour l'échantillon choisi et crée différentes feuilles de calcul ; chaque feuille de calcul présente différentes vues des données en lecture.
- Cette opération peut prendre quelques instants, en fonction de la taille de l'échantillon et du matériel informatique.

IMPORTANT ! En raison des limites de Microsoft Excel, le logiciel de visualisation ne peut pas afficher plus de 1 048 576 lignes par feuille. Seules les 1 048 576 premières lignes de données seront affichées si l'un des fichiers de sortie d'analyse de données pour l'échantillon en cours contient un nombre de lignes supérieur à cette limite.

Si cette limite est atteinte, le programme affiche un message indiquant la ou les feuilles concernées par cette limite. Cliquez sur **OK** pour continuer.

Même si les données ne sont pas toutes affichées, cela n'a pas de conséquences pour les graphiques et ceux-ci seront tout de même corrects.

IMPORTANT ! Les graphiques sont générés en supposant une orientation d'affichage **de gauche à droite**. Dans les pays pour lesquels le sens de lecture habituel est de droite à gauche, les paramètres d'Excel par défaut peuvent être incompatibles, et les graphiques ne sont pas générés correctement. Pour y remédier, l'option d'affichage dans Excel doit être changée en **De gauche à droite**. Cette option peut être configurée en sélectionnant *Office* > *Excel Options* (*Options Excel*) dans Excel 2007 ou *File* (*Fichier*) > *Options* dans Excel 2010/2013. Sélectionnez **Advanced** (**Options avancées**), puis faites défiler vers le bas jusqu'aux paramètres *Display* (*Afficher*). Changez l'option *Default direction* (*Orientation par défaut*) en **Left-to-right** (**De gauche à droite**), puis cliquez sur **OK**. La prochaine fois qu'Excel sera ouvert, les feuilles de calcul liront de gauche à droite et le logiciel de visualisation générera les graphiques correctement.

En cas de problème ou pour toute assistance supplémentaire, veuillez contacter le support technique Invivoscribe au +1 (858) 224 6600, du lundi au vendredi, de 7 h à 17 h heure du Pacifique, ou à l'adresse e-mail support@invivoscribe.com.

8.6.1. Si un message s'affiche concernant l'option *Complex Formatting (Mise en forme complexe)*, cliquez sur **Yes** (**Oui**).

8.7 **Enregistrer le fichier de visualisation.** Lorsque le traitement est terminé, un message apparaît indiquant que le programme a terminé. Il est vivement recommandé d'enregistrer le fichier à ce stade.

IMPORTANT ! Lors de l'enregistrement du fichier, veillez à sélectionner **Save as** (**Enregistrer sous**) et à choisir un classeur Excel (XLSX ou XLS), car l'action par défaut **Save** (**Enregistrer**) ou le fait de cliquer sur l'icône de disquette remplace la macro XLSM par le classeur actuel. Cela nécessitera de réinstaller la macro avec une nouvelle copie provenant du CD du logiciel en vue des analyses ultérieures.

8.7.1. Si un message d'avertissement s'affiche à propos de l'enregistrement d'un classeur sans macro, cliquez sur le bouton **Yes** (**Oui**).

8.8 Le classeur terminé pour Leader, IGH FR1, IGH FR2, IGH FR3, IGK, TRB et TRG comprend les huit feuilles de calcul suivantes décrites dans le Tableau 7.

Tableau 7. Feuilles de calcul incluses dans le fichier de sortie Excel.

Classeur terminé pour Leader, *IGH* FR1, *IGH* FR2, *IGH* FR3, *IGK*, *TRB* et *TRG*	
Read Summary *(Résumé de lecture)*	Contient une liste des 200 séquences uniques les plus importantes, triées par nombre et précisant : ▪ À quel événement de recombinaison V-J chaque séquence correspond le mieux ▪ La séquence ▪ La longueur de la séquence ▪ Le % de lectures totales réel ▪ Le % de lectures totales cumulé sous forme de liste pour toutes les lectures (*TRG* uniquement) ▪ Le % de lectures totales cumulé sous forme de liste pour les 200 premières lectures (*IGH*/*IGK* et *TRB*) ▪ Dans quelle mesure la lecture couvre le gène V identifié ciblé par les amorces (%) (*IGH*/*IGK*) ▪ La séquence putative CDR3 Leader et *IGH* FR1 précisent également : ▪ Le taux de mutations du gène V partiel (%) Leader, *IGH* FR1, *IGH* FR2, *IGH* FR3 précisent également : ▪ Si le réarrangement doit se situer à l'intérieur du cadre (Y/N/n/a [O/N/s.o.]). Si l'utilisation du segment de gène V ou J ne peut pas être prédite en raison de la qualité de l'alignement, le logiciel ne sera pas en mesure de prédire si le gène se situe à l'intérieur du cadre et le résultat sera n/a. Si le réarrangement ne doit pas contenir un codon stop (Y/N [O/N]). Si la position à l'intérieur du cadre ne peut pas être prédite, il n'est pas possible de prédire avec exactitude si un codon stop sera présent. ▪ La séquence putative CDR3
Merged Read Summary *(Résumé de lecture fusionné)*	Même format que la feuille de travail *Read Summary* (Résumé de lecture), mais uniquement pour les 10 premières lectures après fusion avec les 500 premières si les lectures ne se différencient l'une de l'autre que par 1 ou 2 nts.
VJ Sequence Frequency Graph *(Graphique de la fréquence de la séquence VJ)*	Histogramme représentant les 200 fréquences de séquence les plus importantes pour chaque famille V – famille J (pour *IGH*/*IGK* et *TRB*) ou recombinaison de gène V – gène J (pour *TRG*).
VJ Usage *(Utilisation de VJ)*	Comptage brut et fréquences des 200 recombinaison famille V – Famille J (pour *IGH*/*IGK* et *TRB*) ou gene V – Gene J (pour *TRG*) les plus représentés
VJ Usage Percent Graph *(Graphique de pourcentage d'utilisation de VJ)*	Un graphique 3D représentant les 200 fréquences de séquence les plus importantes pour chaque famille V – famille J (pour *IGH/IGK* et *TRB*) ou recombinaison de gène V – gène J (pour *TRG*).
VJ Usage Raw Graph *(Graphique brut d'utilisation de VJ)*	Un graphique 3D représentant les 200 numérations brutes les plus importantes pour chaque famille V – famille J (pour *IGH/IGK* et *TRB*) ou recombinaison de gène V – gène J (pour *TRG*).
VJ Sequence Frequency *(Fréquence de la séquence VJ)*	Une liste « empilée » de toutes les recombinaisons possibles de la famille V – famille J (pour *IGH*/*IGK* et *TRB*) ou du gène V – gène J (pour *TRG*), chacune comportant une liste triée des 200 premières lectures (du % le plus élevé au plus faible).
Unique Reads *(Lectures uniques)*	Fichier FASTA comportant des séquences uniques, triées en fonction de leur nombre.

Remarque : *IGH* figurant dans le tableau ci-dessus fait référence à l'analyse des gènes Leader, *IGH* FR1, *IGH* FR2 et *IGH* FR3. Les graphiques de la fréquence de ces gènes cibles précisent uniquement *IGH* dans le titre. Le dossier de sortie des données indique si la région Leader ou l'une des régions charpentes (FR) a été analysée par le logiciel, et cette information est incluse dans le champ d'emplacement du fichier de données, en haut de la feuille de calcul Read Summary (Résumé de lecture).

Remarque : *l'analyse d'IGK* inclut également des informations sur d'autres éléments du locus *IGK* impliqués dans les recombinaisons (IGKINTR et IGKDEL).

Remarque : *l'analyse de TRB* inclut également des informations sur d'autres éléments du locus *TRB* impliqués dans les recombinaisons de la région D – J.

Remarque : les 200 séquences les plus importantes : lorsque spécifié ci-dessus, les feuilles de calcul affichent les séquences les plus importantes (jusqu'à 200) pour Leader, *IGH* FR1, *IGH* FR2, *IGH* FR3, *TRB* et *IGK*. Toutes les séquences sont affichées pour *TRG* dans le Read Summary (Résumé de lecture).

Remarque : *la séquence putative CDR3* est déterminée par la correspondance d'une définition exacte d'une séquence CDR3. Toute mutation des points d'ancrage ou perturbation du cadre de lecture entraînera une valeur *not found (non trouvé)*.

IMPORTANT ! Les assignations des gènes V/J présentées dans le *Read Summary (Résumé de lecture)* peuvent parfois correspondre à *none (aucune)*. Ces assignations sont basées sur la qualité du meilleur alignement entre la lecture et les gènes de référence. Si le meilleur alignement a un résultat inférieur à un seuil de qualité, l'assignation du gène ne peut pas être déterminée avec certitude et est classée *none (aucune)*.

IMPORTANT ! Le taux de mutation en gène V partiel (%) indiquera *n/a (s.o.)* pour *IGH* FR2 et FR3. La couverture du gène V est limitée en raison des tailles d'amplicon plus courtes de ces cibles, ce qui entraîne des erreurs de calcul du taux de mutation.

- ☐ Lors de l'analyse du statut d'hypermutation somatique d'échantillons, concentrer l'analyse des données sur les 4 premières lectures (limiter les échantillons monoclonaux et biclonaux aux 2 ou 4 premières lectures, respectivement), car le taux d'hypermutation somatique (HMS) et les analyses du cadre interne/du codon stop fournissent davantage d'informations pour ces lectures. En raison de la nature non biaisée des plateformes NGS, de nombreuses lectures au-delà des 4 premières sont séquencées et analysées ; la plupart sont des réarrangements de faible abondance ou peuvent être imputées à des artefacts d'amplification ou de séquençage. Dans ce cas, le taux d'HMS et les analyses du cadre interne/codon stop ne sont pas utiles dans les lectures non clonales. Veuillez noter qu'Invivoscribe utilise GRCh37 pour les séquences de référence, ainsi les résultats peuvent différer légèrement de la base de données IMGT. En outre, un grand nombre de pseudogènes et de gènes non fonctionnels ne sont pas inclus dans l'analyse. Il convient d'interpréter avec prudence les réarrangements dont l'état d'hypermutation somatique est égal à ou proche du seuil de 2 %, en particulier lors de l'exécution du test *IGH* FR1, qui cible uniquement une portion de la région IGHV, car la séquence en amont du site de liaison des amorces ne sera pas évaluée.

9. Support technique et service client

Nous vous sommes reconnaissants de votre fidélité. Nous sommes à votre disposition pour vous aider à comprendre ce logiciel, et notre support technique est disponible du lundi au vendredi pour vous permettre de réaliser les tests efficacement dans votre laboratoire.

Coordonnées

Invivoscribe, Inc
10222 Barnes Canyon Road | Building 1 | San Diego | California 92121-2711 | États-Unis
Téléphone: +1 858 224-6600 | Fax: +1 858 224-6601 | Heures d'ouverture: 7 h – 17 h heure du Pacifique
Service technique: support@invivoscribe.com | Service client: sales@invivoscribe.com | Site internet: www.invivoscribe.com

10. Symboles

Les symboles suivants sont utilisés pour l'étiquetage des produits NGS d'Invivoscribe.

REF	Numéro de référence		Date de péremption
VOL	Volume du réactif	EC REP	Représentant agréé dans la Communauté européenne
LOT	Numéro de lot		Consulter les instructions d'utilisation
	Conditions de conservation	IVD	Destiné au diagnostic *in vitro*
UDI	Identifiant Unique de L'Appareil		Fabricante
UK CA	Conformité Britannique Évaluée	UKRP	Personne responsable au Royaume-Uni
CH REP	Mandataire Suisse	CE	Conformité Européenne
	Application logicielle		

11. Informations légales

Pour des avis légaux concernant ce produit, veuillez consulter l'adresse https://invivoscribe.com/legal-notices/

12. Foire aux questions concernant le logiciel

Q1 : **Pourquoi le Read Summary (Résumé de lecture) de l'analyse *IGH*, *IGK* ou *TRB* comporte-t-il moins de 200 lectures ?**

R1 : Le logiciel LymphoTrack Dx – MiSeq aligne les 200 lectures les plus importantes pendant l'analyse, mais le *Read Summary* (*Résumé de lecture*) contient uniquement les lectures ayant été alignées sur un gène V ou J qui se trouve dans la référence. Si une lecture ne parvient pas à s'aligner sur la référence, elle n'est pas comptabilisée dans le résumé de lecture.

Diverses raisons peuvent expliquer une absence d'alignement sur la référence. Toutefois, le problème se situe généralement au niveau de l'échantillon, conduisant à une grande quantité d'amplification non spécifique. Cet événement peut être dû à une faible numération de cellules cibles dans l'échantillon, à des conditions de PCR incorrectes ou à des types d'échantillons de faible qualité, tels qu'un FFIP fragmenté. Afin de réduire la probabilité de survenue de ces événements, veuillez suivre scrupuleusement les instructions d'utilisation.

Q2 : **Pourquoi l'assignation du gène V ou l'assignation du gène J indique-t-elle *« none »* (aucune) dans le tableau *Read Summary* (*Résumé de lecture*) ?**

R2 : Afin d'assigner une lecture à un gène V et/ou J spécifique, un seuil de qualité minimal doit être atteint pendant l'alignement. Si l'alignement est de mauvaise qualité, une assignation ne peut être effectuée avec certitude et celle-ci comporte la mention « ***none*** » (aucune). Veuillez noter qu'une lecture ne peut être assignée à « none » à la fois pour les gènes V et J, car elle n'apparaîtrait pas dans le rapport, conformément à la condition exposée à la question 2.

Q3 : **Le *VJ Sequence Frequency Graph* (*Graphique de la fréquence de la séquence VJ*) comporte des valeurs étranges sur l'ordonnée. Le *% total reads (% de lectures totales*) est en millions, comment est-ce possible ?**

R3 : Cela est certainement dû à une exécution du programme LymphoTrack Dx avec un type de séparateur décimal incorrect pour votre région. Consultez la section 6 *Exécution du logiciel d'analyse de données*, en particulier en ce qui concerne le choix du séparateur décimal.

Q4 : **Toutes les feuilles de calcul des graphiques sont vides, mais les feuilles de calcul sans graphique comportent des données. Où les graphiques sont-ils passés ?**

R4 : Cela peut être dû à une mauvaise orientation par défaut du programme Excel, à savoir *de droite à gauche* et non *de gauche à droite*. Si les en-têtes des colonnes d'une feuille de calcul vide commencent par la lettre **A** sur la droite de la feuille de calcul, vous devez modifier l'orientation en *De gauche à droite* avant d'utiliser la macro de visualisation des données. Pour modifier ce paramètre, cliquez sur le bouton **Office** dans le coin supérieur gauche, puis sur *Excel Options* (Options Excel), cliquez sur **Advanced** (Options avancées), puis faites défiler vers le bas jusqu'à la section *Display* (Afficher). Une option intitulée *Default direction* (Orientation par défaut) apparaît ici et le fait de sélectionner **Left-to-right** (De gauche à droite) permettra par la suite aux analyses de visualisation de données de fonctionner correctement. Remarque : l'orientation est alors modifiée de façon permanente jusqu'à ce qu'elle soit rétablie. La fermeture d'Excel n'inverse pas cette modification.

Q5 : **Le programme est interrompu après avoir cliqué sur *Launch Program* (*Démarrer le programme*) et le message d'erreur suivant s'affiche :**

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. (L'application s'est terminée avec une erreur : Nom de fichier non compatible. Veuillez utiliser uniquement des fichiers FASTQ générés avec MiSeq MSR. Pour obtenir de l'aide concernant cette erreur, veuillez contacter le support technique Invivoscribe.

R5 : Le logiciel LymphoTrack Dx – MiSeq utilise des fichiers FASTQ à lecture appariée générés par le logiciel MiSeq Reporter (MSR) intégré à l'instrument MiSeq. Si le logiciel MSR n'a pas été utilisé pour générer des fichiers de séquence, la définition des bases doit être effectuée à l'aide du logiciel MSR. Veuillez vous reporter au Guide de l'utilisateur MiSeq pour plus d'informations.

Il est important de définir les bases en utilisant le logiciel MSR car, même si d'autres programmes peuvent générer des résultats très similaires à ceux de MSR (par exemple Bcl2fastq), ces programmes peuvent également générer des résultats très différents selon les paramètres spécifiques utilisés. Le test LymphoTrack Dx Assay est valable uniquement pour l'analyse de fichiers FASTQ générés avec le logiciel MSR.

Q6 : **Pourquoi les résultats de l'analyse semblent-ils inclure un moins grand nombre de lectures que prévu ?**

R6 : S'il y a très peu de lectures totales analysées, voire aucune, d'après le Read Summary (Résumé de lecture), vérifiez que vous avez coché le bon type d'analyse (Leader, *IGH* FR1, *IGH* FR2, *IGH* FR3, *IGK, TRB* ou *TRG*) lors de l'exécution du logiciel LymphoTrack Dx – MiSeq. Si le nombre de lectures totales semble être très réduit, comparez la taille des fichiers FASTQ à analyser avec la taille des fichiers FASTQ présents sur l'instrument MiSeq. Si la taille des fichiers sur l'instrument MiSeq est plus importante, transférez de nouveau les données de séquence afin de bien analyser l'intégralité des données.

L'outil de séquençage MiSeq affiche différents messages lors de la production des données de séquençage. Un message peut indiquer *« Sequencing Complete »* (Séquençage terminé), et un autre peut ensuite indiquer *« Analysis Complete »* (Analyse terminée). Le transfert des données FASTQ entre ces deux messages peut rendre l'analyse LymphoTrack Dx incomplète.

Q7 : **Pourquoi le *Read Summary* (*Résumé de lecture*) comprend-il une lecture, figurant parmi les dix premières, qui n'apparaît pas dans le *Merged Read Summary* (*Résumé de lecture*) ?**

R7 : La lecture introuvable a peut-être été fusionnée avec une lecture d'un rang supérieur, ou il est possible que son alignement soit exceptionnellement médiocre. Lors du processus de fusionnement, les lectures contenues dans le résumé de lecture sont alignées sur celles du fichier de lectures uniques. Cependant, les lectures du résumé de lecture ont été orientées de sorte que l'alignement corresponde à la référence, et il est donc possible qu'elles soient anti-sens par rapport au gène de référence. Cela se produit uniquement lorsque l'alignement est particulièrement mauvais ou qu'un alignement sens ou anti-sens serait presque équivalent, mais que l'anti-sens est légèrement meilleur. Dans ce cas, il est préférable d'ignorer complètement la lecture.

IMPORTANT : **seules les lectures du *Merged Read Summary* (*Résumé de lecture fusionné*) peuvent être utilisées pour mettre en évidence la clonalité.**

Q8 : **Pourquoi l'analyse prend-elle plus de temps que prévu (plus d'une heure par échantillon) ?**

R8 : Consultez le fichier journal pour connaître le temps nécessaire à chaque étape. Sous la ligne *INFO: Decompressing files* (*INFO : Décompression des fichiers*), une autre ligne indique le nombre de fichiers et le temps utilisé pour la décompression. Ce nombre est exprimé en millisecondes. Si la décompression prend plusieurs heures, vérifiez que les fichiers cibles se trouvent sur un serveur de fichiers. En général, l'exécution d'un processus E/S intensif, comme c'est le cas de la décompression, sur un réseau est beaucoup plus longue. Si possible, copiez les fichiers sur un ordinateur local ou décompressez les fichiers directement avant d'exécuter ce logiciel. Veillez à ce que les fichiers compressés et décompressés ne se trouvent pas dans le même dossier. Le logiciel ne poursuivra pas l'opération si les fichiers de type FASTQ et FASTQ GZ se trouvent ensemble dans le dossier cible.

Si les étapes de parsing et d'analyse subséquentes prennent beaucoup de temps, il est possible que vos fichiers FASTQ soient très volumineux. Des fichiers FASTQ volumineux peuvent par exemple être générés lorsque moins de 24 index sont contenus dans une seule analyse MiSeq et que la profondeur de couverture des échantillons séquencés est donc plus importante. Un second exemple est celui d'une bibliothèque déséquilibrée qui contiendrait une proportion plus élevée d'échantillons d'un même index par rapport à la proportion d'échantillons d'autres index. Le temps d'exécution estimé dans ces instructions d'utilisation part du principe que l'analyse comporte 24 index et que les bibliothèques sont équilibrées.

Q9 : **L'analyse Q30 rapportée par l'instrument MiSeq ne répond pas à la spécification Illumina. Si l'analyse LymphoTrack est effectuée et que les contrôles semblent avoir donné les résultats attendus, l'analyse a-t-elle vraiment échoué ?**

R9 : L'analyse Q30 est calculée en temps réel au cours du processus de définition des bases. À ce stade, aucun autre traitement de données n'a été effectué, tel que le démultiplexage et l'élimination des adaptateurs. Autrement dit, toute séquence d'un index de qualité médiocre qui se produit en dehors de la séquence de l'adaptateur est prise en compte dans le score Q30. La séquence en dehors de celle de l'adaptateur est généralement de faible qualité, et ce scénario se produit souvent avec les dimères d'amorces. Si une bibliothèque contient un niveau suffisant de dimères d'amorces, une chute de l'analyse Q30 peut survenir. À titre de confirmation, la sortie LymphoTrack Reporter affiche le score Q30 de l'index tel qu'il est calculé à partir de toutes les lectures éliminées dans cet index. Si le score Q30 d'un index échoue également à la spécification Illumina, l'utilisateur doit considérer que cet index n'est pas conforme.

Q10 : **Pourquoi les 10 principales lectures fusionnées sont-elles différentes lorsque les mêmes données sont analysées à nouveau ?**

R10 : Le nombre de lectures fusionnées dans les 10 principales peut varier légèrement lorsque les données sont analysées plusieurs fois en raison du tri des séquences lorsque plusieurs séquences ont exactement le même nombre de lectures. Dans les cas où plusieurs séquences ont des nombres de lectures identiques, le logiciel peut les trier et les classer de manière aléatoire. Les différences observées devraient être insignifiantes et ne pas affecter l'interprétation des résultats.