

Este documento y su contenido son propiedad de Illumina, Inc. y sus afiliados (“Illumina”) y están previstos solamente para el uso contractual de sus clientes en conexión con el uso de los productos descritos en él y no para ningún otro fin. Este documento y su contenido no se utilizarán ni distribuirán con ningún otro fin ni tampoco se comunicará, divulgará ni reproducirá en ninguna otra forma sin el consentimiento previo por escrito de Illumina. Illumina no transfiere mediante este documento ninguna licencia bajo sus derechos de patente, marca comercial, copyright ni derechos de autor ni similares derechos de terceros.

Para asegurar el uso correcto y seguro de los productos descritos en este documento, el personal cualificado y adecuadamente capacitado debe seguir las instrucciones incluidas en este de manera rigurosa y expresa. Se debe leer y entender completamente todo el contenido de este documento antes de usar estos productos.

SI NO SE LEE COMPLETAMENTE EL DOCUMENTO Y NO SE SIGUEN EXPRESAMENTE TODAS LAS INSTRUCCIONES DESCRITAS EN ESTE, PODRÍAN PRODUCIRSE DAÑOS EN EL PRODUCTO, LESIONES A LAS PERSONAS, INCLUIDOS LOS USUARIOS Y OTROS Y DAÑOS EN OTRA PROPIEDAD.

ILLUMINA NO ASUME RESPONSABILIDAD ALGUNA DERIVADA DEL USO INCORRECTO DE LOS PRODUCTOS AQUÍ DESCRITOS (INCLUIDAS LAS PIEZAS O EL SOFTWARE).

© 2016 Illumina, Inc. Todos los derechos reservados.

Illumina, MiSeq, MiSeqDx, el color naranja calabaza y el diseño de las bases de streaming son marcas comerciales de Illumina, Inc. o sus afiliados en EE. UU. u otros países. Todos los demás nombres, logotipos y marcas comerciales son propiedad de sus respectivos dueños.

Historial de revisiones

N.º de documento	Fecha	Descripción del cambio
N.º de documento 15038356 v02	Noviembre de 2017	Se han actualizado las marcas normativas.
N.º de documento 15038356 v01	Diciembre de 2016	Información corregida en la tabla Targets (Objetivos) para el kit universal 1.0: se ha cambiado inserciones a deleciones en la fila Deletions (Deleciones). Se ha eliminado Chrome de la lista de exploradores compatibles con el uso de MiSeq Reporter fuera del instrumento. Errores de formato corregidos.
N.º de referencia 15038356 Rev. A	Marzo de 2014	Publicación inicial

Índice

Historial de revisiones	iii
Índice	iv
Capítulo 1 Descripción general	1
Introducción	2
Visualización de MiSeq Reporter	3
Conceptos de MiSeq Reporter	4
Interfaz de MiSeq Reporter	5
Puesta de un análisis en cola por segunda vez	13
Métricas de análisis	14
Procedimientos de análisis	16
Carpeta MiSeqAnalysis	17
Capítulo 2 Visualización de datos	19
Introducción	20
Requisitos del archivo de entrada	21
Flujo de trabajo Custom Amplicon (Amplicón personalizado)	22
Archivos de resultados de análisis de los ensayos de FQ	35
Capítulo 3 Instalación y solución de problemas	37
Requisitos de MiSeq Reporter fuera del instrumento	38
Instalación de MiSeq Reporter fuera del instrumento	39
Uso de MiSeq Reporter fuera del instrumento	41
Solución de problemas de MiSeq Reporter	42
Apéndice A Archivos de resultados de análisis del kit universal 1.0	45
Tipos de archivo de resultados de análisis	46
Formato de archivo BAM	47
Formato de archivo VCF	48
Archivo de cobertura de amplicones	51
Archivos de resultados complementarios	52
Índice	53
Asistencia técnica	55

[Página en blanco intencionada]

Descripción general

Introducción	2
Visualización de MiSeq Reporter	3
Conceptos de MiSeq Reporter	4
Interfaz de MiSeq Reporter	5
Puesta de un análisis en cola por segunda vez	13
Métricas de análisis	14
Procedimientos de análisis	16
Carpeta MiSeqAnalysis	17



Introducción

El instrumento MiSeqDx™ incluye tres aplicaciones de software que funcionan en secuencia para generar imágenes de grupos de la celda de flujo, realizar análisis de imágenes y llamadas de bases, y llevar a cabo análisis secundarios integrados en el instrumento.

- ▶ Durante el experimento, MiSeq Operating Software (MOS) captura imágenes de grupos de la celda de flujo para análisis de imágenes, además de hacer funcionar la platina de la celda de flujo, ejecutar los comandos para dispensar reactivos y cambiar las temperaturas de la celda de flujo.
- ▶ El software de análisis primario integrado, análisis en tiempo real (RTA), realizará análisis de imágenes y llamadas de base, y asignará una puntuación de calidad a cada base de cada ciclo a medida que progresa el experimento. La finalización del análisis primario de RTA inicia MiSeq Reporter para que comience el análisis secundario.
- ▶ MiSeq Reporter realiza un análisis secundario integrado en el instrumento de las llamadas de bases y puntuaciones de calidad generadas por RTA durante el experimento de secuenciación. MiSeq Reporter se ejecuta como un servicio de Windows y se puede visualizar a través de un explorador web. Otra posibilidad es instalarlo en un ordenador fuera del instrumento. Para obtener más información, consulte *Instalación de MiSeq Reporter fuera del instrumento* en la página 39.

Acerca de las aplicaciones de servicios de Windows

Las aplicaciones de servicios de Windows realizan funciones específicas sin necesidad de intervención por parte del usuario y continúan ejecutándose en segundo plano siempre que se esté ejecutando Windows. Dado que MiSeq Reporter se ejecuta como un servicio de Windows, este comienza automáticamente el análisis secundario cuando ha finalizado el análisis primario.

Secuenciación durante el análisis

Los recursos informáticos del instrumento MiSeqDx se han concebido para secuenciar o analizar. Si se inicia un nuevo experimento de secuenciación en MiSeqDx antes de que finalice el análisis secundario de un experimento anterior, aparece un cuadro de diálogo de confirmación. Tras confirmar el experimento de secuenciación, se detiene el análisis secundario.

Para reiniciar el análisis, utilice la función **Requeue** (Reposicionar en la cola) en la interfaz de MiSeq Reporter después de finalizar el nuevo experimento de secuenciación. En ese punto, el análisis secundario comenzará desde el principio.

Visualización de MiSeq Reporter

La interfaz de MiSeq Reporter solo se puede visualizar a través de un explorador web. Para ver la interfaz de MiSeq Reporter, abra un explorador web en un ordenador con acceso a la misma red que el instrumento MiSeqDx. Conecte el servicio HTTP en el puerto **8042** mediante uno de los siguientes métodos:

- ▶ Conecte utilizando la dirección IP del instrumento seguida de 8042.

Dirección IP	Puerto de servicio HTTP	Dirección HTTP
10.10.10.10, por ejemplo	8042	10.10.10.10:8042

- ▶ Conecte utilizando el nombre de red de MiSeqDx seguido de 8042.

Nombre de red	Puerto de servicio HTTP	Dirección HTTP
MiSeqDx01, por ejemplo	8042	MiSeqDx01:8042

Para instalaciones de MiSeq Reporter fuera del instrumento, conecte utilizando el método para aplicaciones de servicios instaladas de manera local, **localhost** seguido de 8042.

Fuera del instrumento	Puerto de servicio HTTP	Dirección HTTP
localhost	8042	localhost:8042

Para obtener más información, consulte *Instalación de MiSeq Reporter fuera del instrumento* en la página 39.

Conceptos de MiSeq Reporter

Los conceptos y términos siguientes son comunes en MiSeq Reporter.

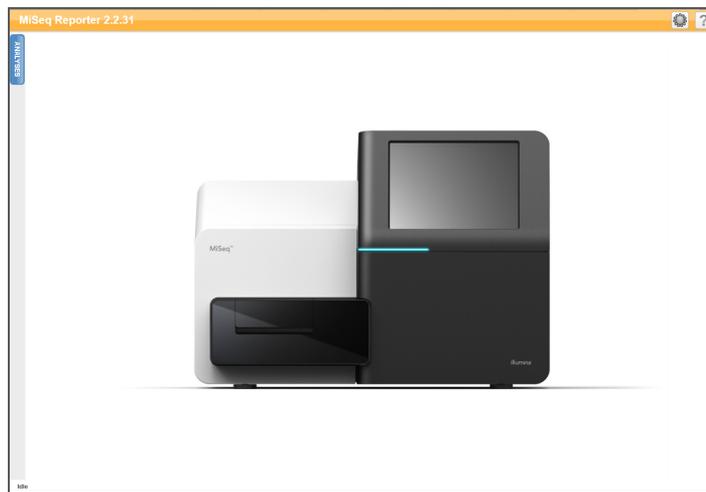
Concepto	Descripción
Manifiesto	Archivo que especifica el genoma de referencia y las regiones de referencia objetivo que se deben utilizar en el paso de alineación. El archivo de manifiesto que se utiliza en los ensayos de fibrosis quística se carga previamente en MiSeqDx.
Repositorio	Una carpeta que alberga los datos generados durante los experimentos de secuenciación. Cada carpeta del experimento es una subcarpeta dentro del repositorio.
Carpeta del experimento	Estructura de carpetas que contiene el software de análisis principal RTA (carpeta MiSeqOutput) o carpeta que contiene MiSeq Reporter (MiSeqAnalysis).
Hoja de muestras	Un archivo de valores separados por coma (*.csv) que contiene la información necesaria para configurar y analizar un experimento de secuenciación, incluida una lista de muestras y sus secuencias de índice. Se crea fuera del instrumento mediante Illumina Worklist Manager. La hoja de muestras debe proporcionarse durante los pasos de configuración del experimento en MiSeqDx. Después de que comience el experimento, la hoja de muestras cambia de nombre automáticamente a SampleSheet.csv y se copia en las carpetas del experimento: MiSeqOutput y MiSeqAnalysis.
Flujo de trabajo	Un procedimiento del análisis secundario que realiza MiSeq Reporter. El flujo de trabajo de cada experimento se especifica en la información de la hoja de muestras.

Interfaz de MiSeq Reporter

Cuando MiSeq Reporter se abre en el explorador, aparece la pantalla principal con una imagen del instrumento en el centro. Los iconos de configuración y ayuda se encuentran en la esquina superior derecha, mientras que la ficha Analyses (Análisis) se encuentra en la esquina superior izquierda.

- ▶ **Ayuda de MiSeq Reporter:** seleccione el icono de ayuda para abrir la documentación de MiSeq Reporter en la ventana del explorador.
- ▶ **Configuración:** seleccione el icono de configuración  para cambiar la ruta del repositorio y la URL del servidor.
- ▶ **Ficha Analyses (Análisis):** seleccione Analyses (Análisis) para ampliar la ficha. La ficha Analyses (Análisis) muestra una lista de experimentos de análisis que se han finalizado, puesto en cola para su análisis o que actualmente se encuentran en proceso.

Figura 1 Pantalla principal de MiSeq Reporter

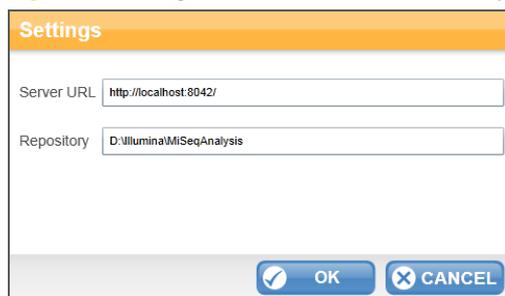


Configuración de URL del servidor o del repositorio

Utilice la función de configuración  para cambiar la URL del servidor y la ruta del repositorio:

- ▶ **Server URL (URL del servidor):** servidor en el que se ejecuta MiSeq Reporter.
- ▶ **Repository path (Ruta del repositorio):** ubicación de la carpeta de análisis en la que se graban los archivos de resultados.

Figura 2 Configuración de URL del servidor y del repositorio



Normalmente, no es necesario que cambie esta configuración a menos que MiSeq Reporter se ejecute fuera del instrumento.

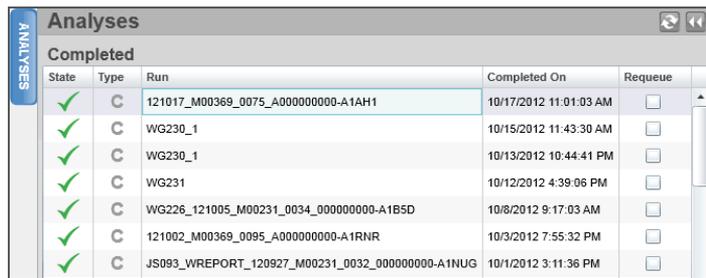
En tal caso, establezca la ruta del repositorio en la ubicación de red de la carpeta MiSeqOutput. Para obtener más información, consulte *Uso de MiSeq Reporter fuera del instrumento* en la página 41.

Ficha Analyses (Análisis)

La ficha Analyses (Análisis) muestra todos los experimentos de secuenciación que se encuentran ubicados en el repositorio especificado. En esta ficha, se pueden abrir los resultados de cualquier experimento que figure en ella o se puede volver a poner en cola un experimento seleccionado para su análisis.

Seleccione el icono de **Refresh Analysis List**  (Actualizar lista de análisis) en la esquina superior derecha para actualizar la lista en cualquier momento.

Figura 3 Ficha Analyses (Análisis) ampliada



State	Type	Run	Completed On	Requeue
✓	C	121017_M00369_0075_A000000000-A1AH1	10/17/2012 11:01:03 AM	<input type="checkbox"/>
✓	C	WG230_1	10/15/2012 11:43:30 AM	<input type="checkbox"/>
✓	C	WG230_1	10/13/2012 10:44:41 PM	<input type="checkbox"/>
✓	C	WG231	10/12/2012 4:39:06 PM	<input type="checkbox"/>
✓	C	WG226_121005_M00231_0034_000000000-A1BSD	10/8/2012 9:17:03 AM	<input type="checkbox"/>
✓	C	121002_M00369_0095_A000000000-A1RNR	10/3/2012 7:55:32 PM	<input type="checkbox"/>
✓	C	JS093_WREPORT_120927_M00231_0032_000000000-A1NUG	10/1/2012 3:11:36 PM	<input type="checkbox"/>

Las columnas de la ficha Analyses (Análisis) son State (Estado), Type (Tipo), Run (Experimento), Completed On (Finalizado el) y Requeue (Volver a poner en cola):

- ▶ **State** (Estado): muestra el estado actual del análisis utilizando uno de los tres iconos de estado.

Tabla 1 Iconos de estado de análisis

Icono	Descripción
	Indica que el análisis secundario se ha realizado correctamente.
	Indica que el análisis secundario está en curso.
	Indica que se han producido errores y que el análisis secundario no se ha realizado correctamente.

- ▶ **Type** (Tipo): muestra el flujo de trabajo de análisis asociado con cada experimento mediante una designación de una sola letra. En el caso de los ensayos de FQ y el kit universal 1.0, la designación de letra es C.
- ▶ **Run** (Experimento): el nombre de la carpeta del experimento en las carpetas MiSeqOutput y MiSeqAnalysis.
- ▶ **Completed On** (Finalizado el): la fecha en que se finalizó el análisis secundario.
- ▶ **Requeue** (Volver a poner en cola): seleccione la casilla de verificación para volver a poner en cola un trabajo específico para su análisis. Aparece el botón Requeue (Volver a poner en cola). Para obtener más información, consulte *Puesta de un análisis en cola por segunda vez* en la página 13.

Cuando se pone en cola el análisis, el experimento aparece en la parte inferior de la ficha Analyses (Análisis) y se indica como si estuviera en curso con el icono .

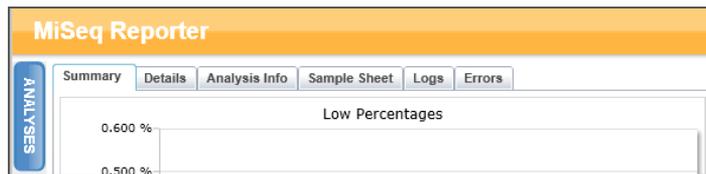
Figura 4 Experimento en cola en la ficha Analyses (Análisis)

✓	C	JS092_120921_M00369_0092_A000000000-A1AHY	9/22/2012 9:20:38 PM	<input type="checkbox"/>
✓	C	WG223	9/20/2012 10:15:33 PM	<input type="checkbox"/>
Queued				
State	Type	Run	Started/Queued On	
⌚	C	121017_M00231_0020_A000000000-A1B5F	10/17/2012 12:11:33 PM	

Fichas Analysis Information and Results (Información de análisis y Resultados)

Después de seleccionar un experimento en la ficha Analyses (Análisis), la información y los resultados de dicho experimento aparecerán en una serie de fichas en la interfaz de MiSeq Reporter: Summary (Resumen), Details (Detalles), Analysis Info (Información del análisis), Sample Sheet (Hoja de muestras), Logs (Registros) y Errors (Errores). La información sobre las fichas Analysis Info (Información del análisis) y Sample Sheet (Hoja de muestras) aparece al principio. Todas las fichas se complementan cuando finaliza el análisis.

Figura 5 Fichas Information (Información) y Results (Resultados)

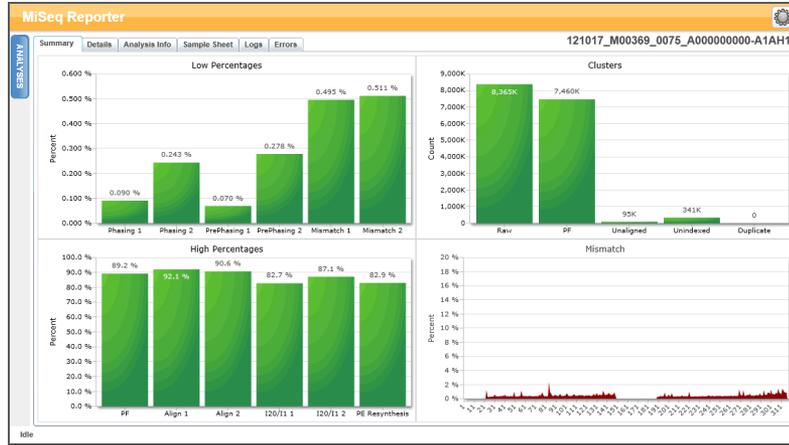


Ficha Summary (Resumen)

La ficha Summary (Resumen) contiene un resumen de los resultados de los análisis. En la ficha Summary (Resumen) aparecen cuatro gráficos:

- ▶ Gráfico **Low Percentages** (Porcentajes bajos): muestra los valores de hebra retrasada, hebra adelantada y las incoherencias en porcentajes. Unos porcentajes bajos indican unas estadísticas del experimento buenas. Para obtener más información, consulte *Hebra retrasada y hebra adelantada* en la página 14.
- ▶ Gráfico **High Percentages** (Porcentajes altos): muestra los grupos que superan el filtro, la alineación con respecto a una referencia y las intensidades en porcentajes. Unos porcentajes altos indican unas estadísticas del experimento buenas.
- ▶ Gráfico **Clusters** (Grupos): muestra el número de grupos sin procesar, los grupos que superan el filtro, los grupos que no se alinearon, los grupos que no se asociaron con un índice y los duplicados.
- ▶ Gráfico **Mismatch** (Incoherencias): muestra las incoherencias por ciclo. Una incoherencia se refiere a cualquier incoherencia entre la lectura de secuenciación y un genoma de referencia tras la alineación.

Figura 6 Ficha Summary (Resumen)



Ficha Details (Detalles)

La ficha Details (Detalles) contiene detalles de los resultados del análisis. Las tablas y gráficos siguientes pueden aparecer en la ficha Details (Detalles) en función del ensayo o kit utilizado:

- ▶ Tabla **Samples** (Muestras): resume los resultados de secuenciación de cada muestra.
- ▶ Tabla **Targets** (Objetivos): muestra las estadísticas de las regiones objetivo de una muestra seleccionada. (Solamente kit universal 1.0)
- ▶ Tabla **Variants** (Variantes): muestra las diferencias entre el ADN de muestra y la referencia.
- ▶ Gráfico **Coverage** (Cobertura): muestra el nivel de profundidad empleado en la secuenciación de la muestra mediante la medición del número de bases presentes en la secuencia de la muestra por cada posición de la referencia.
- ▶ Gráfico **Qscore** (Puntuación de calidad): muestra la puntuación de calidad media, que es la probabilidad estimada de que se produzca un error de llamada de bases. Para obtener más información, consulte *Gráfico Qscore (Puntuación de calidad)* en la página 33.
- ▶ Gráfico **Variant Score** (Puntuación de variantes): muestra la ubicación de las variantes de nucleótido simple (SNV) y las inserciones y deleciones.

Figura 7 Ejemplo de ficha Details (Detalles) del ensayo de 139 variantes de FQ

The screenshot shows the MiSeq Reporter Details page for sample 121017_M00369_0075_A00000000-A1AH1. It displays two main sections:

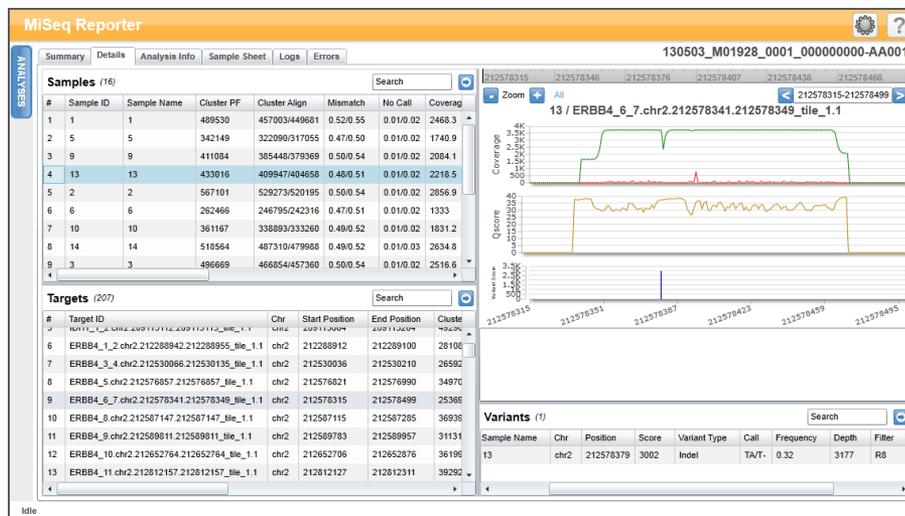
Samples (46)

#	Sample ID	Sample Name	Call Rate	Performance	Control	Comment
1	NA01445	NA01445	100.00	Pass		
2	NA02828	NA02828	100.00	Pass		
3	NA04330	NA04330	100.00	Pass		
4	NA07339	NA07339	100.00	Pass		
5	NA07381	NA07381	100.00	Pass		
6	NA07441	NA07441	100.00	Pass		
7	NA07732	NA07732	100.00	Pass		
8	NA07854	NA07854	100.00	Pass		
9	NA07857	NA07857	100.00	Pass		

Variants (3)

#	Sample ID	Sample Name	Mutation Name	Type	dbSNP rsID	CFTR Gene Region	Genomic Location	cDNA Name (HGVS)	Protein Name (HGVS)	Result	Col
1	NA01445	NA01445	PolyTG/PolyT	PolyTG/PolyT	N/A	Intron 9	117188661-117188689	c.1210-12T[S_S]	N/A	(TG)10(T)/7(TG)10(T)9	S3
2	NA01445	NA01445	F508del	DEL	rs113993960	Exon 11	117199845	c.1521_1523delCTT	p.Phe508del	HET	301
3	NA01445	NA01445	W1282X	SNV	rs77010898	Exon 23	117282620	c.3846G>A	p.Tyr1282X	HET	301

Figura 8 Ejemplo de ficha Details (Detalles) para el kit universal 1.0



Los resultados de las tablas Samples (Muestras), Targets (Objetivos) o Variants (Variantes) se pueden exportar de manera individual a un archivo de texto con el icono de **exportación de datos de tabla a archivo de texto**. Esta exportación no modifica el archivo del informe del análisis.



En el caso de los ensayos de FQ, los resultados se pueden exportar al archivo del informe del análisis de FQ con el icono de **exportación de datos a informe de FQ**. Para obtener más información, consulte *Archivos de resultados de análisis de los ensayos de FQ* en la página 35.

Ficha Analysis Info (Información del análisis)

La ficha Analysis Info (Información de análisis) contiene información logística sobre el experimento y el análisis.

Figura 9 Ficha Analysis Info (Información del análisis)

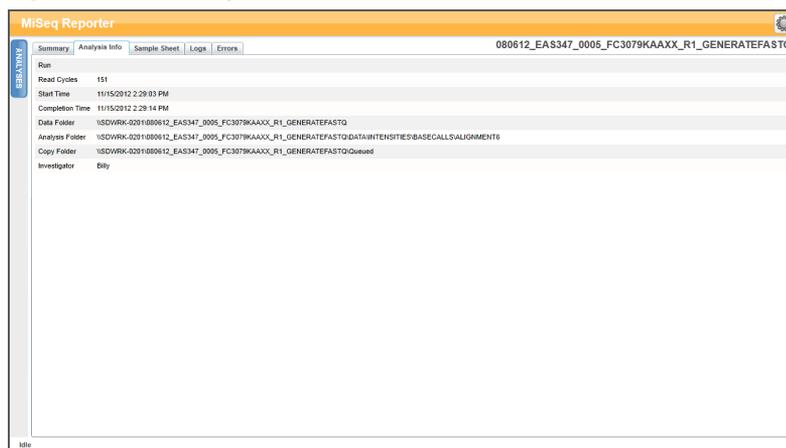


Tabla 2 Ficha Analysis Info (Información del análisis)

Fila	Descripción
Read Cycles (Ciclos de lectura)	Una representación del número de ciclos en cada lectura, incluida la anotación de cualquier lectura del índice. Por ejemplo, un experimento anotado como 151, 8 (I), 8 (I) significa lo siguiente: 151 indica 151 ciclos para la primera lectura, 8 ciclos para la primera lectura del índice, 8 ciclos para la segunda lectura del índice y una lectura final de 151 ciclos.
Start Time (Hora de inicio)	La hora del reloj a la que se inició el análisis secundario.
Completion Time (Hora de finalización)	La hora a la que finalizó el análisis secundario.
Data Folder (Carpeta de datos)	El nivel raíz de la carpeta de resultados producido por el software de análisis primario de RTA (MiSeqOutput), que contiene todos los resultados del análisis primario y secundario del experimento.
Analysis Folder (Carpeta de análisis)	La ruta completa a la carpeta de alineación de la carpeta MiSeqAnalysis (Data\Intensities\BaseCalls\Alignment).
Copy Folder (Carpeta de copia)	La ruta completa hasta la subcarpeta de muestras en cola en la carpeta MiSeqAnalysis.

Ficha Sample Sheet (Hoja de muestras)

La ficha Sample Sheet (Hoja de muestras) contiene parámetros del experimento especificados en la hoja de muestras, además de ofrecer herramientas para editar esta última y, a continuación, volver a poner en cola el experimento.

Figura 10 Ficha de hoja de muestras, ejemplo de kit universal 1.0

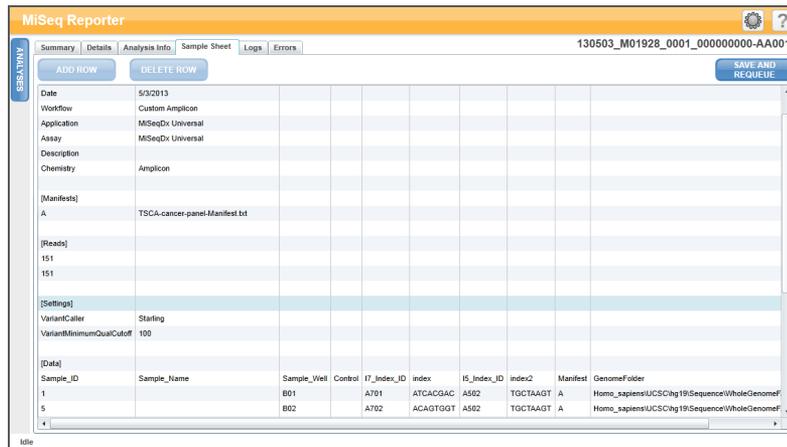


Tabla 3 Contenido de la ficha Sample Sheet (Hoja de muestras)

Fila	Descripción
Date (Fecha)	La fecha en que se realizó el experimento de secuenciación.

Tabla 3 Contenido de la ficha Sample Sheet (Hoja de muestras)

Fila	Descripción
Workflow (Flujo de trabajo)	El flujo de trabajo de análisis del experimento. En el caso de los ensayos de FQ y el kit universal 1.0, el nombre del flujo de trabajo es Custom Amplicon (Amplicón personalizado).
Application (Aplicación)	Nombre de la aplicación. Lo utiliza el software Illumina Worklist Manager. Este campo indica el ensayo o el kit que se utiliza para el experimento.
Assay (Ensayo)	Nombre del ensayo o el kit.
Chemistry (Composición química)	El nombre químico identifica los fragmentos de fórmula que se utilizan para la fórmula del experimento específico. En los experimentos de MiSeqDx, el nombre químico es "amplicon" (amplicón).
Manifests (Manifiestos)	El nombre del archivo de manifiesto que especifica el genoma de referencia y las regiones de referencia objetivo que se deben utilizar en el paso de alineación.
Reads (Lecturas)	El número de ciclos llevados a cabo en la Lectura 1 y la Lectura 2. Las lecturas del índice no se incluyen en esta sección.
Settings (Configuración)	Parámetros del experimento opcionales.
Data (Datos)	El ID de la muestra, el nombre de la muestra, las secuencias de índices y la ruta a la carpeta de genoma. Los requisitos varían en función del flujo de trabajo.

Ficha Logs (Registros)

La ficha Logs (Registros) muestra cada paso que se ha llevado a cabo durante el análisis. Estos pasos se registran en archivos de registro ubicados en la carpeta Logs (Registros). Se genera un resumen en AnalysisLog.txt, que es un archivo importante para solucionar posibles problemas.

Ficha Errors (Errores)

La ficha Errors (Errores) muestra cualquier error que se haya producido durante el análisis. Se genera un resumen en AnalysisError.txt, que es un archivo importante para solucionar posibles problemas.

Edición de la hoja de muestras en MiSeq Reporter

Los datos de la hoja de muestras se pueden editar para un experimento específico en la ficha Sample Sheet (Hoja de muestras) en la interfaz web de MiSeq Reporter. Para editar la hoja de muestras se necesitan un ratón y un teclado.



PRECAUCIÓN

La edición de la información de la hoja de muestras debe realizarse extremando el cuidado y la precaución. El seguimiento de las muestras podría verse alterado y traducirse en unos resultados incorrectos.

- Para editar una fila de la hoja de muestras, haga clic en el campo y, a continuación, realice los cambios necesarios.

- ▶ Para añadir una fila a la hoja de muestras, haga clic en la fila y seleccione **Add Row** (Añadir fila). La nueva fila aparecerá debajo de la fila seleccionada.

ADD ROW

- ▶ Para eliminar una fila de la hoja de muestras, haga clic en la fila y seleccione **Delete Row** (Eliminar fila).

DELETE ROW

- ▶ Tras finalizar los cambios en la hoja de muestras, seleccione **Save and Requeue** (Guardar y volver a poner en cola). De esta forma, se guardarán los cambios y se iniciará el análisis secundario con la hoja de muestras editada.

SAVE AND
REQUEUE

- ▶ Si se realizó un cambio en la hoja de muestras por error, haga clic en la ficha adyacente antes de guardar los cambios. Aparecerá un mensaje de advertencia que indica que no se han guardado los cambios. Haga clic en **Discard** (Descartar) para deshacer los cambios.

DISCARD

Almacenamiento de gráficos como imágenes

MiSeq Reporter ofrece la posibilidad de guardar una imagen de los gráficos que genera un experimento. Haga clic con el botón derecho sobre cualquier ubicación en la ficha Summary (Resumen) o en la ubicación de los gráficos en la ficha Details (Detalles) y, a continuación, haga clic con el botón izquierdo en **Save Image As** (Guardar imagen como). Cuando se le solicite, asigne un nombre al archivo y vaya a una ubicación para guardar dicho archivo.

Todas las imágenes se guardan en formato JPG. Los gráficos se exportan como un solo gráfico de todos los que figuran en la ficha. Para utilizar esta opción es necesario un ratón.

Puesta de un análisis en cola por segunda vez

Se puede volver a poner en cola un análisis en la interfaz web de MiSeq Reporter. Antes de continuar, compruebe que no haya ningún experimento de secuenciación en curso.

Cada vez que vuelve a poner en cola un análisis, se crea una nueva carpeta de alineación en la carpeta MiSeqAnalysis con un número secuencial junto al nombre. Por ejemplo, Alignment, Alignment1, Alignment2.

```
MiSeqAnalysis \<NombreCarpetaExperimento> \Data \Intensities \
BaseCalls \Alignment2
```

- 1 En la interfaz web de MiSeq Reporter, haga clic en **Analyses** (Análisis).
- 2 Localice el experimento en la lista de experimentos disponibles y haga clic en la casilla de verificación Requeue (Volver a poner en cola) junto al nombre del experimento. Si no figura el experimento, cambie el repositorio especificado a la ubicación correcta. Para obtener más información, consulte *Configuración de URL del servidor o del repositorio* en la página 5.

Figura 11 Puesta de un análisis en cola por segunda vez



Analyses					
Completed					REQUEUE
State	Type	Run	Completed On	Requeue	
✓	C	121017_M00369_0075_A000000000-A1AH1	10/17/2012 11:01:03 AM	<input checked="" type="checkbox"/>	
✓	C	WG230_1	10/15/2012 11:43:30 AM	<input type="checkbox"/>	
✓	C	WG230_1	10/13/2012 10:44:41 PM	<input type="checkbox"/>	
✓	C	WG231	10/12/2012 4:39:06 PM	<input type="checkbox"/>	
✓	C	WG226_121005_M00231_0034_000000000-A1B5D	10/8/2012 9:17:03 AM	<input type="checkbox"/>	
✓	C	121002_M00369_0095_A000000000-A1RNR	10/3/2012 7:55:32 PM	<input type="checkbox"/>	

- 3 Haga clic en **Requeue** (Volver a poner en cola). El icono de la columna State (Estado) a la izquierda del nombre del experimento cambia para mostrar que ese análisis se encuentra en curso .



NOTA

Si el análisis no se inicia, asegúrese de que están presentes los siguientes archivos de entrada en la carpeta del experimento del análisis: SampleSheet.csv, RTACComplete.txt y RunInfo.xml.

Métricas de análisis

Durante el experimento de secuenciación, el análisis en tiempo real (RTA) genera archivos de datos que incluyen métricas de análisis que utiliza MiSeq Reporter para el análisis secundario. Las métricas que aparecen en los informes de los análisis secundarios son grupos que superan el filtro, puntuaciones de calidad de las llamadas de bases y valores de hebra retrasada y hebra adelantada.

Grupos que superan el filtro

Los grupos que superan el filtro constituyen una medición de la calidad de los grupos. Este filtro elimina los datos menos fiables mediante la filtración de incidencias para eliminar cualquier lectura que no cumpla los requisitos de calidad general. Los grupos que superan el filtro se representan con PF en los informes de análisis.

Puntuaciones de calidad

Una puntuación de calidad, o puntuación Q, es una predicción de la probabilidad de obtener una llamada de bases incorrecta. Durante el experimento de secuenciación, se registran las puntuaciones de calidad de la llamada de bases de cada ciclo. Durante el análisis, las puntuaciones de calidad se registran en los archivos FASTQ en un formato codificado ASCII.

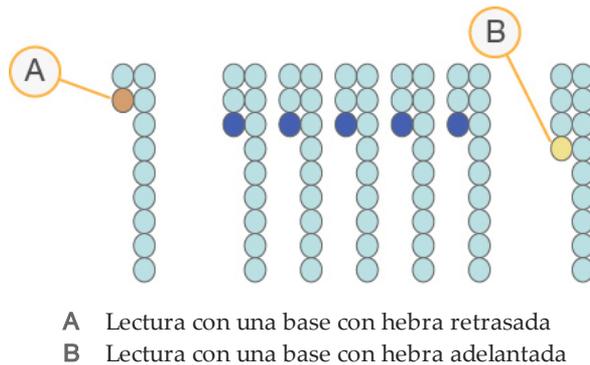
En la siguiente tabla, figura la relación entre la puntuación de calidad y la probabilidad de error.

Puntuación de calidad	Probabilidad de error
Q40	0,0001 (1 entre 10 000)
Q30	0,001 (1 entre 1000)
Q20	0,01 (1 entre 100)
Q10	0,1 (1 entre 10)

Hebra retrasada y hebra adelantada

Durante la reacción de secuenciación, cada cadena de ADN de un grupo se amplía en una base por cada ciclo. Una pequeña porción de las cadenas puede encontrarse fuera de la hebra con el ciclo de incorporación actual, ya sea porque se encuentre una base atrás (hebra retrasada) o se adelante una base (hebra adelantada). Los valores de hebra retrasada y hebra adelantada indican una estimación de la fracción de moléculas que se han dividido o predividido en cada ciclo.

Figura 12 Hebra retrasada y hebra adelantada



El número de ciclos realizados en una lectura es de un ciclo más que el número de ciclos analizados. Por ejemplo, un experimento de 150 ciclos de "paired-end" realiza dos lecturas de 151 ciclos (2×151) para un total de 302 ciclos. Al final del experimento, se habrán analizado 2×150 ciclos. Se requiere realizar el ciclo adicional en la lectura 1 y la lectura 2 para los cálculos de hebra adelantada.

Procedimientos de análisis

MiSeq Reporter realiza un análisis secundario con una serie de procedimientos de análisis, que incluyen el demultiplexado, la generación de archivos FASTQ, la alineación y las llamadas de variantes.

Demultiplexado

El demultiplexado es el primer paso en el análisis si la hoja de muestras contiene varias muestras y el experimento tiene lecturas del índice.

El demultiplexado separa datos de las muestras agrupadas en función de las secuencias de índice cortas que marcan las muestras de diferentes bibliotecas. Cada secuencia de lectura del índice se compara con las secuencias de índice especificadas en la hoja de muestras. En este paso, no se considera ningún valor de calidad.

Generación de archivos FASTQ

Tras el demultiplexado, este procedimiento genera archivos intermedios en formato de archivo FASTQ, que es un formato de texto empleado para representar secuencias. Los archivos FASTQ son la entrada principal para el paso de alineación. Los archivos FASTQ contienen las lecturas de cada muestra y las puntuaciones de calidad, excepto las lecturas de los grupos que no hayan superado el filtro.

Alineación

La alineación compara secuencias con la referencia para identificar una relación entre las secuencias y asigna una puntuación en función de las regiones de similitud. Las lecturas alineadas se escriben en archivos con formato BAM.

Para los datos generados en MiSeq Reporter, MiSeqDx utiliza un algoritmo de Smith-Waterman de bandas que realiza las alineaciones de secuencias locales para determinar regiones similares entre dos secuencias. En lugar de considerar la secuencia total, el algoritmo de Smith-Waterman compara los segmentos de todas las longitudes posibles. Las alineaciones locales resultan útiles para secuencias no similares de las que se sospeche que contengan regiones de similitud dentro de una secuencia de mayor longitud.

Llamadas de variantes

Las llamadas de variantes registran los polimorfismos de nucleótido único (SNP), las inserciones y deleciones, y otras variantes estructurales.

Para los datos generados en el instrumento MiSeqDx, las llamadas de variantes las realiza el llamador de variantes Starling de MiSeq Reporter. Este llamador llama a los SNP y a las inserciones y deleciones pequeñas, y resume la profundidad y la probabilidad de los errores de cada sitio del genoma. Para cada llamada de inserciones y deleciones o SNP, la probabilidad de un error se proporciona como puntuación de calidad de una variante.

Tras la finalización, Starling genera informes en formato HTML de SNP e inserciones y deleciones, así como archivos de texto delimitado por tabulaciones con variantes en formato de llamada de variantes (VCF). Para obtener más información, consulte *Formato de archivo VCF* en la página 48.

Carpeta MiSeqAnalysis

La carpeta MiSeqAnalysis es la carpeta del experimento principal de MiSeq Reporter. La relación entre las carpetas de experimento MiSeqOutput y MiSeqAnalysis se resume de la siguiente manera:

- ▶ Durante la secuenciación, el análisis en tiempo real (RTA) completa la carpeta MiSeqOutput con los archivos que se han generado durante el análisis primario.
- ▶ Excepto en el caso de imágenes de enfoque o en miniatura, las copias del RTA se archivan en la carpeta MiSeqAnalysis en tiempo real. Cuando finaliza el análisis principal, el RTA escribe el archivo RTAComplete.xml en las dos carpetas del experimento.
- ▶ MiSeq Reporter supervisa la carpeta MiSeqAnalysis y comienza el análisis secundario cuando el archivo RTAComplete.xml aparece.
- ▶ A medida que continúa el análisis secundario, MiSeq Reporter escribe los archivos de resultados de análisis en la carpeta MiSeqAnalysis y, a continuación, copia los archivos en la carpeta MiSeqOutput.

[Página en blanco intencionada]

Visualización de datos

Introducción	20
Requisitos del archivo de entrada	21
Flujo de trabajo Custom Amplicon (Amplicón personalizado)	22
Archivos de resultados de análisis de los ensayos de FQ	35



Introducción

MiSeq Reporter realiza el análisis secundario y genera diversos tipos de información específicos del ensayo tras realizar el análisis. Los resultados aparecen en la interfaz web de MiSeq Reporter en forma de gráficos y tablas para cada experimento. Los productos de MiSeqDx incluyen los enumerados en la tabla siguiente:

Producto	Descripción
Ensayo de 139 variantes de fibrosis quística	Detecta 139 variantes clínicamente relevantes del gen CFTR de un máximo de 48 muestras.
Ensayo de secuenciación clínica de fibrosis quística	Detecta mutaciones en las regiones de codificación de proteínas incluidas las fronteras intrón/exón, dos deleciones de gran tamaño y dos mutaciones intrónicas profundas en el gen CFTR de un máximo de 8 muestras.
Kit universal 1.0	Conjunto de reactivos y consumibles utilizado junto con oligonucleótidos personalizados proporcionados por el usuario para llevar a cabo la resecuenciación selectiva de regiones genómicas de interés específicas.

Requisitos del archivo de entrada

MiSeq Reporter requiere los siguientes archivos de análisis principal generados durante el experimento de secuenciación para llevar a cabo el análisis secundario o volver a poner en cola el análisis. Los archivos del análisis primario, tales como *.bcl, *.filter y *.locs, son necesarios para efectuar el análisis.

No es necesario mover ni copiar archivos en otra ubicación antes de iniciar el análisis. Los archivos necesarios se copian automáticamente en la carpeta MiSeqAnalysis durante el proceso de secuenciación.

Nombre de archivo	Descripción
RTAComplete.txt	Un archivo de marcador que indica que el procesamiento de RTA ha finalizado. La presencia de este archivo activa MiSeq Reporter para que ponga en cola el análisis.
SampleSheet.csv	Proporciona parámetros para el experimento y los análisis posteriores. Al inicio del experimento, la hoja de muestras se copia en el nivel raíz de la carpeta del experimento y cambia su nombre a SampleSheet.csv.
RunInfo.xml	Contiene información del experimento de alto nivel, como el número de lecturas y ciclos en el experimento de secuenciación, y si se ha indexado o no una lectura.

Genomas y bases de datos instaladas previamente

MiSeqDx incluye bases de datos y genomas instalados previamente.

Preinstalado	Descripción
Bases de datos	dbSNP para humanos, versión 131 refGene para humanos
Genomas	humano (<i>Homo sapiens</i>) versión hg19

Flujo de trabajo Custom Amplicon (Amplicón personalizado)

El flujo de trabajo Custom Amplicon (Amplicón personalizado) empleado para los ensayos de FQ y el kit universal 1.0 evalúan las regiones breves de ADN amplificado, o los amplicones, para las variantes. La secuenciación de amplicones enfocada permite una alta cobertura de regiones particulares a lo largo de un gran número de muestras.

Después del demultiplexado y de la generación del archivo FASTQ, el flujo de trabajo realiza los siguientes pasos:

- ▶ **Alineación:** los grupos de cada muestra se alinean en relación con las secuencias de amplicones especificadas en el archivo de manifiesto.
 - En el caso de los datos "paired-end", cada lectura se evalúa inicialmente en términos de alineación respecto a las secuencias de sonda significativas de dicha lectura. La Lectura 1 se evalúa comparándola con el complemento inverso de los oligonucleótidos específicos del locus descendente (DLSO) y la Lectura 2 se evalúa comparándola con los oligonucleótidos específicos del locus ascendente (ULSO). Si el comienzo de una secuencia de lectura coincide con una secuencia de sonda con solo una incoherencia, la longitud completa de la lectura se alineará con la secuencia objetivo de amplicones de dicha secuencia de sonda. Esta alineación se realiza a lo largo de la longitud de las secuencias objetivo de amplicones mediante una alineación de Smith-Waterman en bandas.
 - Las inserciones y deleciones dentro de los DLSO y los ULSO no se observan debido a la química del ensayo.
- ▶ **Evaluación "paired-end":** en el caso de experimentos "paired-end", se considera la alineación de mayor puntuación de cada lectura. Si ninguna de las lecturas se alineó o se alineó a cromosomas diferentes, estas lecturas se marcan como un par sin resolver. Además, si las dos alineaciones proceden de amplicones diferentes (esto es, filas diferentes en la sección de objetivos del manifiesto), las lecturas se marcarán como un par sin resolver.
- ▶ **Agrupación/clasificación:** las lecturas se agrupan por muestra y cromosoma y, a continuación, se clasifican según la posición del cromosoma. Los resultados se incluyen en un archivo BAM por muestra.
- ▶ **Llamada de variantes:** las mutaciones se identifican mediante el llamador de variantes. Para obtener más información, consulte *Llamadas de variantes* en la página 16.
- ▶ **Análisis y anotación de variantes:** con una base de datos SNP (dbsnp.txt) previamente instalada, se marca cualquier mutación conocida en el archivo del informe del análisis.
- ▶ **Generación de informes de estadísticas:** las estadísticas se resumen y se incluyen en un informe.

Ficha Summary (Resumen)

La información que aparece en la ficha Summary (Resumen) incluye un gráfico de bajos porcentajes, un gráfico de altos porcentajes, un gráfico de grupos y un gráfico de incoherencias.

Figura 13 Ejemplo de ficha Summary (Resumen)



Gráfico Low Percentages (Porcentajes bajos)

Eje Y	Eje X	Descripción
Percent (Porcentaje)	Phasing 1 (Hebra retrasada 1)	El porcentaje de moléculas en un grupo que se encuentra detrás del ciclo actual en la Lectura 1.
	Phasing 2 (Hebra retrasada 2)	El porcentaje de moléculas en un grupo que se encuentra detrás del ciclo actual en la Lectura 2.
	PrePhasing 1 (Hebra adelantada 1)	El porcentaje de moléculas en un grupo que se procesó antes del ciclo actual en la Lectura 1.
	PrePhasing 2 (Hebra adelantada 2)	El porcentaje de moléculas en un grupo que se procesó antes del ciclo actual en la Lectura 2.
	Mismatch 1 (Incoherencia 1)	El porcentaje promedio de incoherencias de la Lectura 1 de todos los ciclos.
	Mismatch 2 (Incoherencia 2)	El porcentaje promedio de incoherencias de la Lectura 2 de todos los ciclos.

Gráfico High Percentages (Porcentajes altos)

Eje Y	Eje X	Descripción
Percent (Porcentaje)	PF	El porcentaje de grupos que superan el filtro.
	Align 1 (Alineación 1)	El porcentaje de grupos que se alineó con la referencia en la Lectura 1.
	Align 2 (Alineación 2)	El porcentaje de grupos que se alineó con la referencia en la Lectura 2.
	I20 / I1 1	El índice de intensidades en el ciclo 20 en relación con las intensidades en el ciclo 1 de la Lectura 1.
	I20 / I1 2	El índice de intensidades en el ciclo 20 en relación con las intensidades en el ciclo 1 de la Lectura 2.
	PE Resynthesis (Resíntesis de PE)	El índice de intensidades del primer ciclo de la Lectura 1 en relación con las intensidades del primer ciclo de la Lectura 2.

Gráfico Clusters (Grupos)

Eje Y	Eje X	Descripción
Clusters (Grupos)	Raw (Sin procesar)	El número total de grupos detectados en el experimento.
	PF	El número total de grupos que superan el filtro en el experimento.
	Unaligned (No alineados)	El número total de grupos que superan el filtro que no se alineó con el genoma de referencia, si procede. Los grupos que no se han indexado no se incluyen en el recuento de no alineados.
	Unindexed (No indexados)	El número total de grupos que superan el filtro que no se asociaron con ninguna secuencia de índice en el experimento.
	Duplicate (Duplicados)	Este valor no se aplica a los ensayos de FQ o el kit universal 1.0, por lo que siempre será cero.

Gráfico Mismatch (Incoherencias)

Eje Y	Eje X	Descripción
Percent (Porcentaje)	Cycle (Ciclo)	Muestra el porcentaje de incoherencias de todos los grupos de un experimento por ciclo.

Ficha Details (Detalles) para el ensayo de 139 variantes de FQ

La información que aparece en la ficha Details (Detalles) del ensayo de 139 variantes de FQ incluye una tabla de muestras y una tabla de variantes.

Figura 14 Ejemplo de ficha Details (Detalles) del ensayo de 139 variantes de FQ

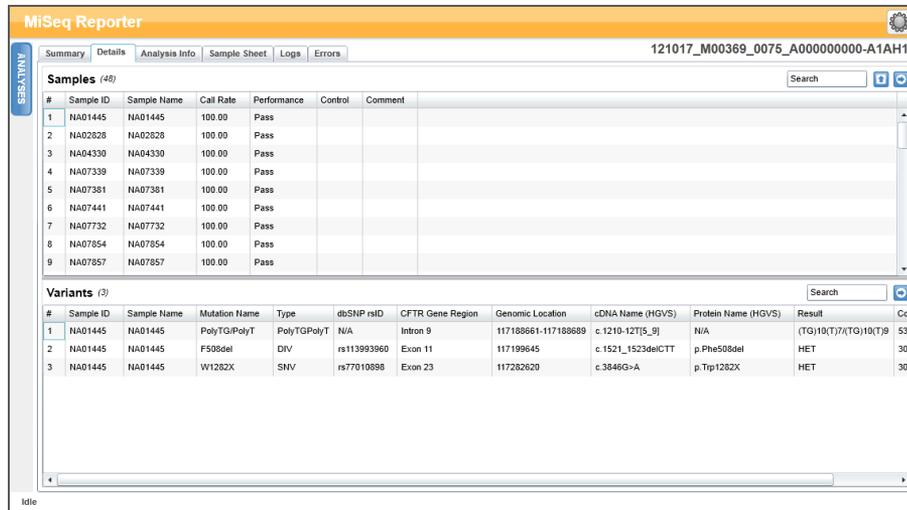


Tabla Samples (Muestras) para el ensayo de 139 variantes de FQ

Columna	Descripción
#	Un número de identificación ordinal en la tabla.

Columna	Descripción
Sample ID (ID de la muestra)	El ID de la muestra de la hoja de muestras. El ID de la muestra debe ser siempre un valor único.
Sample Name (Nombre de la muestra)	El nombre de la muestra de la hoja de muestras.
Call Rate (Índice de llamada)	El número de posiciones mutacionales que cumple un umbral de valor de confianza predefinido dividido entre el total de posiciones mutacionales interrogadas. El índice de llamada se describe por muestra y se incluye en el informe en forma de porcentaje que se calcula como 1 menos [número de posiciones con llamadas incompletas dividido entre el número total de posiciones secuenciadas].
Performance (Rendimiento)	Correcto o incorrecto basándose en el índice de llamada. Para una muestra de control positivo: <ul style="list-style-type: none"> • PASS (APTO): con un índice de llamada $\geq 99\%$ • FAIL (NO APTO): con un índice de llamada $< 99\%$ Para una muestra de control negativo: <ul style="list-style-type: none"> • PASS (APTO): con un índice de llamada $\leq 10\%$ • FAIL (NO APTO): con un índice de llamada $> 10\%$ Para una muestra que no se ha etiquetado como un control positivo o negativo: <ul style="list-style-type: none"> • PASS (APTO): con un índice de llamada $\geq 99\%$ • FAIL (NO APTO): con un índice de llamada $< 99\%$
Control	El tipo de control tal como figura en la hoja de muestras. Los valores son positivos o negativos. Un campo vacío indica la presencia únicamente de la muestra.
Comment (Comentario)	Un campo de texto opcional para comentarios. Los comentarios introducidos en este campo se guardan en el archivo del informe del análisis, MiSeqDxCF139VariantAssay.txt. Si el análisis se vuelve a poner en cola, se creará un nuevo archivo de informe. Los comentarios de un experimento de análisis anterior no pasan al siguiente experimento de análisis.

Tabla Variants (Variantes) para el ensayo de 139 variantes de FQ

Columna	Descripción
#	Un número de identificación ordinal en la tabla.
Sample ID (ID de la muestra)	El ID de la muestra de la hoja de muestras. El ID de la muestra debe ser siempre un valor único.
Sample Name (Nombre de la muestra)	El nombre de la muestra de la hoja de muestras.

Columna	Descripción
Mutations (Common Name) (Mutaciones [Nombre común])	Nombre común de la variante de fibrosis quística, tal como se describe en la base de datos de CFTR2.
Mutation Type (Tipo de mutación)	El tipo de variante: <ul style="list-style-type: none"> • SNV: variante de nucleótido único • DIV: variante de delección/inserción • DEL: delección de gran tamaño • PolyTGPolyT: genotipo PolyTG/PolyT del gen de FQ
dbSNP rsID (rsID de dbSNP)	El rsID de dbSNP de la variante, si procede.
CFTR Gene Region (Región del gen de CFTR)	Región del gen de CFTR (n.º de exón o de intrón) en el que está presente la variante.
Genomic Location (Ubicación genómica)	Ubicación genómica de la variante.
cDNA Name (Nombre de ADNc) (HGVS)	Descripción de una variante en el nivel de ADN que utiliza la nomenclatura de secuencia de ADN de codificación (ADNc) conforme a las recomendaciones de la Sociedad de Variaciones del Genoma Humano (HGVS).
Protein Name (Nombre de proteína) (HGVS)	Descripción de una variante en el nivel de proteínas que utiliza la nomenclatura de secuencia de proteínas conforme a las recomendaciones de la Sociedad de Variación del Genoma Humano (HGVS).
Resultado	Genotipo de variante. Para SNV, DIV y DEL: <ul style="list-style-type: none"> • HET: heterocigota • HOM: homocigota Para la variante PolyTGPolyT, figura el genotipo real. NOTA: PolyTGPolyT figura únicamente cuando se detecta la variante R117H.

Ficha Details (Detalles) para el ensayo de secuenciación clínica de FQ

La información que aparece en la ficha Details (Detalles) del ensayo de secuenciación clínica de FQ incluye una tabla de muestras, una tabla de variantes, un gráfico de cobertura, un gráfico de Qscore y un gráfico de puntuación de variantes.

Figura 15 Ejemplo de ficha Details (Detalles) para el ensayo de secuenciación clínica de FQ

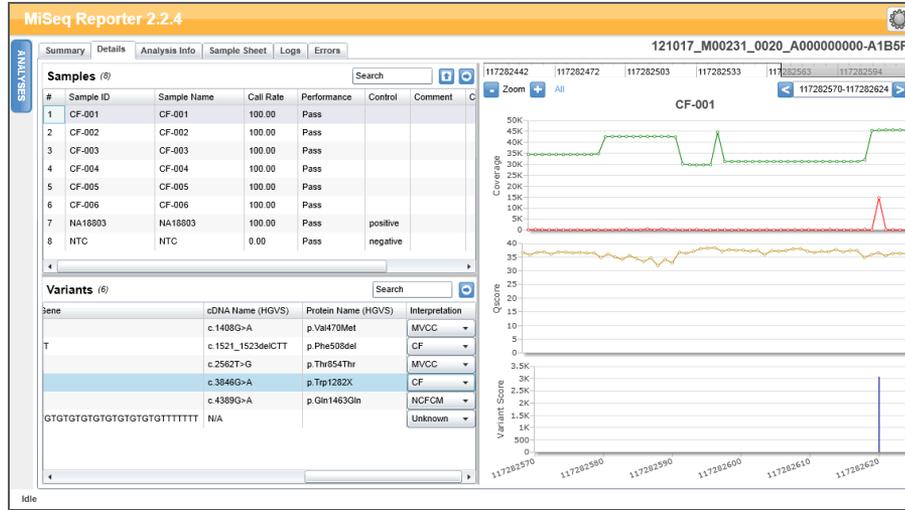


Tabla Samples (Muestras) para el ensayo de secuenciación clínica de FQ

Columna	Descripción
#	Un número de identificación ordinal en la tabla.
Sample ID (ID de la muestra)	El ID de la muestra de la hoja de muestras. El ID de la muestra debe ser siempre un valor único.
Sample Name (Nombre de la muestra)	El nombre de la muestra de la hoja de muestras.
Call Rate (Índice de llamada)	El número de bases que cumple un umbral de puntuación de calidad dividido entre el total de bases interrogadas. El índice de llamada se describe por muestra y se incluye en el informe en forma de porcentaje que se calcula como 1 menos [número de posiciones con llamadas incompletas dividido entre el número total de bases/posiciones secuenciadas].
Performance (Rendimiento)	Correcto o incorrecto basándose en el índice de llamada. Para una muestra de control positivo: <ul style="list-style-type: none"> • PASS (APTO): con un índice de llamada ≥ 99 % • FAIL (NO APTO): con un índice de llamada < 99 % Para una muestra de control negativo: <ul style="list-style-type: none"> • PASS (APTO): con un índice de llamada ≤ 10 % • FAIL (NO APTO): con un índice de llamada > 10 % Para una muestra que no se ha etiquetado como un control positivo o negativo: <ul style="list-style-type: none"> • PASS (APTO): con un índice de llamada ≥ 99 % • FAIL (NO APTO): con un índice de llamada < 99 %

Columna	Descripción
Control	El tipo de control tal como figura en la hoja de muestras. Los valores son positivos o negativos. Un campo vacío indica la presencia únicamente de la muestra.
Comment (Comentario)	Un campo de texto opcional para comentarios. Los comentarios introducidos en este campo se guardan en el archivo del informe del análisis MiSeqDxCFClinicalSequencing.txt. Si el análisis se vuelve a poner en cola, se creará un nuevo archivo de informe. Los comentarios de un experimento de análisis anterior no pasan al siguiente experimento de análisis.
Coordinates Not Called (Coordenadas no llamadas)	Coordenadas del genoma que se encuentran dentro de la región de interés en la que no se registró ninguna llamada debido a unos bajos valores de confianza.

Tabla Variants (Variantes) para el ensayo de secuenciación clínica de FQ

Columna	Descripción
#	Un número de identificación ordinal en la tabla.
Sample ID (ID de la muestra)	El ID de la muestra de la hoja de muestras. El ID de la muestra debe ser siempre un valor único.
Sample Name (Nombre de la muestra)	El nombre de la muestra de la hoja de muestras.
Chr (Cromosoma)	El nombre del cromosoma o el objetivo de referencia.
Position (Posición)	La posición en la que se encontró la variante.
Variant Type (Tipo de variante)	El tipo de variante: <ul style="list-style-type: none"> • SNV: variante de nucleótido único • DIV: variante de deleción/inserción • DEL: deleción de gran tamaño • PolyTGPolyT: genotipo PolyTG/PolyT del gen de FQ
Call (Llamada)	Una cadena que representa cómo la base o las bases cambiaron a esta ubicación en la referencia.
Frequency (Frecuencia)	La fracción de lecturas de la muestra que incluye la variante. Por ejemplo, si la base de referencia en una posición determinada es A y la muestra 1 tiene 60 lecturas A y 40 lecturas T, la variante SNV tiene una frecuencia de variante de 0,4.
Depth (Profundidad)	El número de lecturas de una muestra que cubre una posición determinada.
Filter (Filtro)	Los criterios para una variante filtrada.

Columna	Descripción
dbSNP ID (ID de dbSNP)	El nombre de dbSNP de la variante.
RefGene	El gen conforme a RefGene en el que aparece esta variante.
cDNA Name (Nombre de ADNc) (HGVS)	Descripción de una variante en el nivel de ADN que utiliza la nomenclatura de secuencia de ADN de codificación (ADNc) conforme a las recomendaciones de la Sociedad de Variaciones del Genoma Humano (HGVS).
Protein Name (Nombre de proteína) (HGVS)	Descripción de una variante en el nivel de proteínas que utiliza la nomenclatura de secuencia de proteínas conforme a las recomendaciones de la Sociedad de Variación del Genoma Humano (HGVS).
Interpretación	<p>Este campo permite al médico genetista proporcionar la interpretación clínica de la mutación de cada muestra. Las siguientes opciones se incluyen en la lista desplegable de cada muestra:</p> <ul style="list-style-type: none"> • CF (FQ): que causa FQ. • MVCC: mutación de consecuencia clínica variable. • MOUS: mutación de relevancia desconocida. • NCFCM: no FQ que causa mutación • Desconocido <p>Se puede generar un nuevo informe con el icono.</p>

Columna de interpretación de la tabla de variantes

La columna de interpretación ofrece opciones de selección que permiten al médico genetista proporcionar la interpretación clínica de la mutación de cada muestra. Las siguientes opciones se incluyen en la lista desplegable Interpretation (Interpretación):

- **CF (FQ):** que causa FQ.
- **MVCC:** mutación de consecuencia clínica variable.
- **MOUS:** mutación de relevancia desconocida.
- **NCFCM:** no FQ que causa mutación
- **Desconocido**

Figura 16 Columna de interpretación



Los resultados en las tablas de variantes se pueden exportar de manera individual a un archivo de texto con el icono **Export table data to text file** (Exportar datos de tabla a archivo de texto). Esta exportación no modifica el archivo del informe del análisis.



Después de que el médico genetista haya terminado de determinar la relevancia de la variante, se pueden guardar los ajustes de interpretación en el informe del análisis. Al nombre del archivo del informe del análisis original se le incluirá de manera automática la fecha y la hora.

Gráfico Coverage (Cobertura) para el ensayo de secuenciación clínica de FQ

Eje Y	Eje X	Descripción
Coverage (Cobertura)	Position (Posición)	La curva verde es el número de lecturas alineadas que cubren cada posición en la referencia. La curva roja es el número de lecturas alineadas con una ausencia de llamada en la posición en la referencia. Las SNV y las demás variantes se muestran en forma de picos en la curva roja.

Gráfico Qscore (Puntuación de calidad)

Eje Y	Eje X	Descripción
Qscore (Puntuación de calidad)	Position (Posición)	La puntuación de calidad promedio de las bases en una posición determinada de la referencia.

Gráfico Variant Score (Puntuación de variantes) para el ensayo de secuenciación clínica de FQ

Eje Y	Eje X	Descripción
Score (Puntuación)	Position (Posición)	Muestra gráficamente la puntuación de calidad y la posición de las SNV y de las inserciones y deleciones.

Ficha Details (Detalles) para el kit universal 1.0

La información que aparece en la ficha Details (Detalles) para el kit universal 1.0 incluye una tabla de muestras, una tabla de objetivos, un gráfico de cobertura, un gráfico de puntuación de calidad, un gráfico de puntuación de variantes y una tabla de variantes.

Figura 17 Ejemplo de ficha Details (Detalles) para el kit universal 1.0

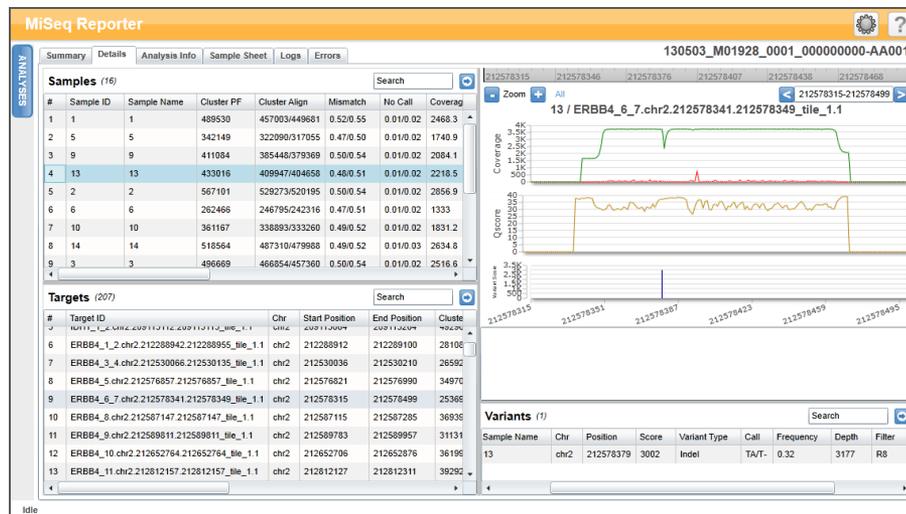


Tabla Samples (Muestras) para el kit universal 1.0

Columna	Descripción
#	Un número de identificación ordinal en la tabla.
Sample ID (ID de la muestra)	El ID de la muestra de la hoja de muestras. El ID de la muestra debe ser siempre un valor único.
Sample Name (Nombre de la muestra)	El nombre de la muestra de la hoja de muestras.
Cluster PF (PF de grupo)	El número de grupos que superan el filtro de cada muestra.
Cluster Align (Alineación de grupos)	El recuento total de grupos que superan el filtro que se alinean para la muestra (lectura 1/lectura 2).
Mismatch (Incoherencia)	El porcentaje de incoherencia respecto de la referencia promediado teniendo en cuenta los ciclos por lectura (lectura 1/lectura 2).
No Call (Ausencia de llamada)	El porcentaje de bases que no se han podido llamar (ausencia de llamada) para la muestra promediado teniendo en cuenta los ciclos por lectura (lectura 1/lectura 2).
Coverage (Cobertura)	Cobertura mediana (número de bases alineadas con una posición de referencia dada) promediada teniendo en cuenta todas las posiciones.
Het SNPs (SNP het)	El número de SNP heterocigóticos detectado para la muestra.
Hom SNPs (SNP hom)	El número de SNP homocigóticos detectado para la muestra.
Insertions (Inserciones)	El número de inserciones detectado para la muestra.
Deletions (Deleciones)	El número de deleciones detectado para la muestra.
Manifest (Manifiesto)	Archivo que especifica el genoma de referencia y las regiones de referencia objetivo que se deben utilizar en el paso de alineación.
Genome (Genoma)	El nombre del genoma de referencia.

Tabla Targets (Objetivos) para el kit universal 1.0

Columna	Descripción
#	Un número de identificación ordinal en la tabla.

Columna	Descripción
Target ID (ID de objetivo)	El nombre del objetivo en el manifiesto.
Chr (Cromosoma)	El nombre del cromosoma o el objetivo de referencia.
Start Position (Posición de inicio)	La posición de inicio de la región objetivo.
End Position (Posición final)	La posición final de la región objetivo.
Cluster PF (PF de grupo)	Número de grupos que superan el filtro del objetivo mostrado por lectura (lectura 1/lectura 2).
Mismatch (Incoherencia)	El porcentaje de bases incoherentes en relación con el objetivo promediado teniendo en cuenta todos los ciclos por lectura. Incoherencia = [media (recuento de errores en ciclos)/grupos que superan el filtro] * 100.
No Call (Ausencia de llamada)	El porcentaje de bases con ausencia de llamada para el objetivo promediado teniendo en cuenta los ciclos por lectura.
Het SNPs (SNP het)	El número de SNP heterocigóticos detectado para el objetivo de todas las muestras.
Hom SNPs (SNP hom)	El número de SNP homocigóticos detectado para el objetivo de todas las muestras.
Insertions (Inserciones)	El número de inserciones detectado para el objetivo de todas las muestras.
Deletions (Deleciones)	El número de deleciones detectado para el objetivo de todas las muestras.
Manifest (Manifiesto)	Archivo que especifica el genoma de referencia y las regiones de referencia objetivo que se deben utilizar en el paso de alineación.

Gráfico Coverage (Cobertura) para el kit universal 1.0

Eje Y	Eje X	Descripción
Coverage (Cobertura)	Position (Posición)	La curva verde es el número de lecturas alineadas que cubren cada posición en la referencia. La curva roja es el número de lecturas alineadas con una ausencia de llamada en la posición en la referencia. Los SNP y las demás variantes se muestran en forma de picos en la curva roja.

Gráfico Qscore (Puntuación de calidad)

Eje Y	Eje X	Descripción
Qscore (Puntuación de calidad)	Position (Posición)	La puntuación de calidad promedio de las bases en una posición determinada de la referencia.

Gráfico Variant Score (Puntuación de variantes) para el kit universal 1.0

Eje Y	Eje X	Descripción
Score (Puntuación)	Position (Posición)	Describe gráficamente la puntuación de calidad de la variante y la posición de los SNP y de las inserciones y las deleciones.

Tabla Variants (Variantes) para el kit universal 1.0

Columna	Descripción
#	Un número de identificación ordinal en la tabla.
Sample ID (ID de la muestra)	El ID de la muestra de la hoja de muestras. El ID de la muestra debe ser siempre un valor único.
Sample Name (Nombre de la muestra)	El nombre de la muestra de la hoja de muestras.
Chr (Cromosoma)	El nombre del cromosoma o el objetivo de referencia.
Position (Posición)	La posición en la que se encontró la variante.
Score (Puntuación)	La puntuación de calidad de la variante correspondiente a esta variante.
Variant Type (Tipo de variante)	El tipo de variante, que puede ser SNP o Indel (inserción y deleción).
Call (Llamada)	Una representación del modo en que la base o las bases han cambiado en esta posición de la referencia. <ul style="list-style-type: none"> • Los SNP se muestran con el formato Referencia > AleloA/AleloB. • Las inserciones se muestran con el formato Referencia/Inserción. G-/GA representa la inserción de A. • Las deleciones se muestran con el formato Referencia/Deleción. AGG/A-- representa la deleción de GG.
Frequency (Frecuencia)	La fracción de lecturas de la muestra que incluye la variante. Por ejemplo, si la base de referencia es A y la muestra 1 tiene 60 lecturas A y 40 lecturas T, el SNP tendrá una frecuencia de variante de 0,4.
Depth (Profundidad)	El número de lecturas de una muestra que cubre una posición determinada.

Columna	Descripción
Filter (Filtro)	Los criterios para una variante filtrada. Si se superan todos los filtros, aparece PASS (APTO) en la columna de filtrado. Para obtener más información, consulte <i>Anotaciones y encabezados de archivos VCF</i> en la página 49.
dbSNP	El nombre de dbSNP de la variante, si procede.
RefGene	El gen conforme a RefGene en el que aparece esta variante.

Archivos de resultados de análisis de los ensayos de FQ

Los resultados de los análisis de los ensayos de FQ se muestran en la ficha Details (Detalles).

Figura 18 Ejemplo de ficha Details (Detalles) del ensayo de 139 variantes de FQ

#	Sample ID	Sample Name	Call Rate	Performance	Control	Comment
1	NA01445	NA01445	100.00	Pass		
2	NA02828	NA02828	100.00	Pass		
3	NA04330	NA04330	100.00	Pass		
4	NA07339	NA07339	100.00	Pass		
5	NA07381	NA07381	100.00	Pass		
6	NA07441	NA07441	100.00	Pass		
7	NA07732	NA07732	100.00	Pass		
8	NA07854	NA07854	100.00	Pass		
9	NA07857	NA07857	100.00	Pass		

#	Sample ID	Sample Name	Mutation Name	Type	dbSNP rsID	CFTR Gene Region	Genomic Location	cDNA Name (HGVS)	Protein Name (HGVS)	Result	Col
1	NA01445	NA01445	PolyTG/PolyT	Poly/TG/PolyT	N/A	Intron 9	117188661-117188689	c.1210-12T[S_9]	N/A	(TG)10(T)7(TG)10(T)9	53
2	NA01445	NA01445	F508del	DEL	rs113893980	Exon 11	117199645	c.1521_1523delCTT	p.Phe508del	HET	307
3	NA01445	NA01445	W1282X	SNV	rs77010898	Exon 23	117282820	c.3846G>A	p.Tip1282X	HET	307



Los resultados en las tablas de variantes se pueden exportar de manera individual a un archivo de texto con el icono **Export table data to text file** (Exportar datos de tabla a archivo de texto). Esta exportación no modifica el archivo del informe del análisis.



Después de que el médico genetista haya terminado de determinar la relevancia de la variante, se pueden guardar los ajustes de interpretación en el informe del análisis. Al nombre del archivo del informe del análisis original se le incluirá de manera automática la fecha y la hora.

Los archivos de resultados de los ensayos de FQ también se resumen en un archivo de texto delimitado por tabulaciones cuyo nombre es el del ensayo utilizado para el experimento. Estos resultados son los mismos que los que se muestran en la ficha Details (Detalles).

- ▶ En el caso del ensayo de 139 variantes de FQ, el archivo se llama `MiSeqDxCf139VariantAssay.txt`.
- ▶ En el caso del ensayo de secuenciación clínica de FQ, el archivo se llama `MiSeqDxCfClinicalSequencingAssay.txt`.

Cuando finaliza el análisis, el archivo de resultados se guarda en la carpeta Alignment (Alineación) del experimento. Por ejemplo:

`MiSeqAnalysis \<NombreCarpetaExperimento> \Data \Intensities \BaseCalls \Alignment`

Si el análisis se ha repetido o se ha vuelto a poner en cola, se escribe un nuevo archivo de informe en la carpeta de alineación del experimento del análisis. Para obtener más información, consulte *Puesta de un análisis en cola por segunda vez* en la página 13.

El archivo de resultados contiene un encabezado que incluye la siguiente información sobre el experimento:

Encabezado	Descripción
Test (Prueba)	Describe la prueba que se ha realizado.
Run ID (ID del experimento)	Se trata del ID del experimento que generó MOS al comienzo del experimento de secuenciación.
Run Date (Fecha del experimento)	Se trata de la fecha, con formato DDMMAA, en la que se inició el experimento de secuenciación en MOS.
Analysis Version (Versión del análisis)	Se trata de la versión de MiSeq Reporter que se utilizó para el análisis.

Figura 19 Ejemplo de encabezado del archivo de resultados del ensayo de 139 variantes de FQ

```

Test CF 139-Variant Assay
For In Vitro Diagnostic Use.
Run ID 140212_M01018_0071_0000000000-A2618
Run Date 140212
Analysis Version 2.2.31.1
    
```

Tras el encabezado, se muestra una sección de resumen de los ID de cada muestra que incluye columnas para cada valor registrado. Para obtener información sobre las descripciones de las columnas, consulte *Ficha Details (Detalles) para el ensayo de 139 variantes de FQ* en la página 24 y *Ficha Details (Detalles) para el ensayo de secuenciación clínica de FQ* en la página 26.



NOTA

El proceso de análisis que genera archivos de resultados no es idéntico entre ensayos de FQ y el kit universal 1.0. Los archivos de resultados generados para el kit universal 1.0 son archivos *.bam, *.vcf y AmpliconCoverage_M#.tsv. Para obtener más información sobre los archivos de resultados para el kit universal 1.0, consulte el Apéndice A Archivos de resultados de análisis del kit universal 1.0.

Instalación y solución de problemas

Requisitos de MiSeq Reporter fuera del instrumento	38
Instalación de MiSeq Reporter fuera del instrumento	39
Uso de MiSeq Reporter fuera del instrumento	41
Solución de problemas de MiSeq Reporter	42



Requisitos de MiSeq Reporter fuera del instrumento

La instalación de una copia de MiSeq Reporter en un ordenador Windows fuera del instrumento permite el análisis secundario de los datos de secuenciación mientras que el instrumento MiSeqDx realiza el siguiente experimento de secuenciación.

Para obtener más información, consulte *Instalación de MiSeq Reporter fuera del instrumento* en la página 39.

Requisitos informáticos

MiSeq Reporter precisa de los siguientes componentes informáticos:

- ▶ Sistema operativo Windows de 64 bits (Vista, Windows 7 y Windows Server 2008 de 64 bits)
- ▶ ≥ 8 GB de RAM como mínimo; se recomienda ≥16 GB de RAM
- ▶ ≥ 1 TB de espacio en disco
- ▶ Procesador de cuatro núcleos (mínimo 2,8 GHz)
- ▶ Microsoft .NET 4

Exploradores admitidos

MiSeq Reporter se puede visualizar con los siguientes exploradores web:

- ▶ Safari 5.1.7 o posterior
- ▶ Firefox 13.0.1 o posterior
- ▶ Internet Explorer 8 o posterior

Descarga y licencias

- 1 Descargue una segunda copia del software MiSeq Reporter en el sitio web de Illumina. Se requiere iniciar una sesión de MyIllumina.
- 2 Acepte el acuerdo del usuario final (EULA) cuando se le solicite durante la instalación. No se requiere ninguna clave de licencia, ya que esta copia adicional se incluye de forma gratuita.

Instalación de MiSeq Reporter fuera del instrumento

Para instalar MiSeq Reporter en un ordenador Windows fuera del instrumento, configure en primer lugar el permiso **Iniciar sesión como servicio** y, a continuación, ejecute el asistente de instalación. A continuación, configure el software para que se dirija a las rutas de repositorio y de GenomePath correctas.

Configuración de cuentas de usuario o grupo en Windows 7

Los derechos de administrador son necesarios para configurar cuentas de usuario o grupo con el objeto de activar el permiso **Iniciar sesión como servicio**. Si es necesario, póngase en contacto con el administrador de las instalaciones.

- 1 En el menú **Inicio** de Windows, seleccione **Panel de control** y, a continuación, **Sistema y seguridad**.
- 2 Haga clic en **Herramientas administrativas** y, a continuación, haga doble clic en **Directiva de seguridad local**.
- 3 En el árbol de configuración de seguridad a la izquierda, haga doble clic en **Directivas locales** y, a continuación, haga clic en **Asignaciones de derechos de usuario**.
- 4 En el panel de detalles de la derecha, haga doble clic en **Iniciar sesión como un servicio**.
- 5 En el cuadro de diálogo Propiedades, haga clic en **Agregar usuario o grupo**.
- 6 Introduzca el nombre de la cuenta de usuario o grupo para este ordenador. Haga clic en **Comprobar nombres** para validar la cuenta.
- 7 Haga clic en **Aceptar** en los cuadros de diálogo abiertos y, a continuación, cierre el panel de control.

Para obtener más información, consulte [technet.microsoft.com/en-us/library/cc739424\(WS.10\).aspx](http://technet.microsoft.com/en-us/library/cc739424(WS.10).aspx) en el sitio web de Microsoft.

Ejecución del asistente de instalación de MiSeq Reporter

- 1 Descargue y descomprima el paquete de instalación de MiSeq Reporter del sitio web de Illumina.
- 2 Haga doble clic en el archivo setup.exe.
- 3 Haga clic en **Next** (Siguiendo) en las indicaciones del asistente de instalación.
- 4 Cuando se le solicite, especifique el nombre de usuario y la contraseña de una cuenta con permiso **Iniciar como servicio**, según se haya configurado en el paso anterior.
- 5 Continúe con las instrucciones restantes.

Configuración de MiSeq Reporter

Para configurar MiSeq Reporter con el fin de localizar la carpeta del experimento y la carpeta del genoma de referencia, edite el archivo de configuración en un editor de texto, como Bloc de notas.

- 1 Navegue hasta la carpeta de instalación (C:\Illumina\MiSeq Reporter, ruta predeterminada) y abra el archivo MiSeq Reporter.exe.config en un editor de texto.

- 2 Localice la etiqueta **Repository** (Repositorio) y cambie **value** (valor) a la ubicación de datos predeterminada en el ordenador fuera del instrumento.

Ejemplo:

```
<add key="Repository" value="E:\Data\Repository" />
```

De manera alternativa, esta ubicación puede ser una ubicación de red a la que puede acceder desde el ordenador fuera del instrumento.

- 3 Localice la etiqueta **GenomePath** y cambie **value** (valor) a la ubicación de la carpeta que contiene los archivos de los genomas de referencia en el formato FASTA.

Ejemplo:

```
<add key="GenomePath" value="E:\MyGenomes\FASTA" />
```

Inicio del servicio de MiSeq Reporter

Tras finalizar la instalación, el servicio de MiSeq Reporter se inicia automáticamente. Si el servicio no se inicia, inícielo de forma manual con las siguientes instrucciones o reinicie el ordenador.

- 1 En el menú **Inicio** de Windows, haga clic con el botón derecho en **Equipo** y seleccione **Administrar**.
- 2 En el árbol de administración del equipo a la izquierda, haga doble clic en **Servicios y aplicaciones** y, a continuación, haga clic en **Servicios**.
- 3 Haga clic con el botón derecho en **MiSeq Reporter** y seleccione **Propiedades**.
- 4 En la ficha General, asegúrese de que el parámetro **Tipo de inicio** esté configurado en **Automático** y, a continuación, haga clic en **Iniciar**.
- 5 En la ficha Iniciar sesión, defina el **nombre de usuario** y la **contraseña** para una cuenta de servicios que tenga permisos de lectura y escritura en el servidor. Illumina recomienda la cuenta **Sistema local** para la mayoría de usuarios. Para obtener asistencia o los requisitos de red específicos del centro, póngase en contacto con el administrador de su centro.
- 6 Haga clic en **Aceptar** en los cuadros de diálogo abiertos y, a continuación, cierre la ventana Administración de equipos.
- 7 Tras iniciar el servicio de MiSeq Reporter, conéctese al software de manera local introduciendo localhost:8042 en un explorador web.

Uso de MiSeq Reporter fuera del instrumento

Para utilizar MiSeq Reporter fuera del instrumento, debe poder accederse a las carpetas que contienen los datos del experimento y los genomas de referencia.

- 1 A menos que se utilice una ubicación de red para los datos de secuenciación y los genomas de referencia, copie las siguientes carpetas en el ordenador local:
 - Copie los datos del experimento del ordenador de MiSeqDx en D:\MiSeqOutput\ - Copie los genomas de referencia del ordenador de MiSeqDx en C:\Illumina\MiSeq Reporter\Genomes.
- 2 Abra un explorador web en <http://localhost:8042>, que abre la interfaz web de MiSeq Reporter.
- 3 Cambie la ruta al repositorio con el icono de **configuración**  en la esquina superior derecha de la interfaz web.



NOTA

La especificación de la ruta del repositorio en la configuración es temporal. La próxima vez que se reinicie el ordenador, la ruta se establecerá de manera predeterminada en la ubicación de repositorio especificada en MiSeq Reporter.exe.config.

- 4 Seleccione **Analyses** (Análisis) en el lado izquierdo de la interfaz web para ver los experimentos disponibles en la ubicación de repositorio especificada.
- 5 Antes de poder volver a poner en cola el análisis de un experimento con la instalación fuera del instrumento de MiSeq Reporter, debe actualizarse la ruta a GenomeFolder en la hoja de muestras. Para ello, vaya a la ficha Sample Sheet (Hoja de muestras). Una vez actualizada la ruta a GenomeFolder, haga clic en **Save and Requeue** (Guardar y volver a poner en cola).

Para obtener más información, consulte *Edición de la hoja de muestras en MiSeq Reporter* en la página 11.

Solución de problemas de MiSeq Reporter

MiSeq Reporter se ejecuta como una aplicación de servicio de Windows. Las cuentas de usuario deben estar configuradas para activar los permisos **Iniciar sesión como servicio** antes de instalar MiSeq Reporter. Para obtener más información, consulte *Configuración de cuentas de usuario o grupo en Windows 7* en la página 39.

Para obtener más información, consulte msdn.microsoft.com/en-us/library/ms189964.aspx.

El servicio no se inicia

Si el servicio no se inicia, compruebe el registro de eventos de Windows y vea los detalles del mensaje de error.

- 1 Abra el **Panel de control** y seleccione **Herramientas administrativas**.
- 2 Seleccione **Visor de eventos**.
- 3 En la ventana del visor de eventos, seleccione **Registros de Windows | Aplicación**. El error que figura en el registro de eventos describe cualquier error de sintaxis en MiSeq Reporter.exe.config. Una sintaxis incorrecta en el archivo MiSeq Reporter.exe.config puede provocar el fallo del servicio.

Los archivos no se copian

Si los archivos no se copian en la ubicación prevista, compruebe los siguientes ajustes:

- 1 Compruebe la ruta a la carpeta de repositorio o la carpeta MiSeqOutput especificada:
 - Para instalaciones fuera del instrumento, compruebe la ubicación del repositorio con el icono de configuración  de la interfaz web de MiSeq Reporter.
 - Para instalaciones integradas en el instrumento, compruebe la ubicación de la carpeta MiSeqOutput en la pantalla Run Options (Opciones de experimento) de MOS en la ficha Folder Settings (Configuración de carpeta).
Debe utilizarse la ruta UNC completa (por ejemplo, \\server1\Runs). Dado que MiSeq Reporter se ejecuta como servicio de Windows, no se reconocen las unidades asignadas por el usuario (por ejemplo, Z:\Runs).
- 2 Confirme el acceso a escritura en la ubicación de la carpeta de resultados. Para obtener asistencia, póngase en contacto con el administrador del centro.
- 3 Compruebe que la función de copia no esté inhabilitada en el archivo MiSeq Reporter.exe.config. Este ajuste se encuentra en la sección <appSettings> y el valor debe estar establecido en **1**.

```
<add key="CopyToRTAOutputPath" value="1"/>
```

Visualización de archivos de registro de un experimento con errores

La visualización de los archivos de registro puede ayudar a identificar errores específicos para la solución de posibles problemas.

- 1 Para ver los archivos de registro con la interfaz del explorador web de MiSeq Reporter, seleccione el experimento en la ficha Analyses (Análisis).
- 2 Seleccione la ficha Logs (Registros) para ver una lista de cada paso que se ha producido durante el análisis. La información del registro se guarda en AnalysisLog.txt, que se encuentra en el nivel raíz de la carpeta MiSeqAnalysis.
- 3 Seleccione la ficha Errors (Errores) para ver una lista de los errores que se han producido durante el análisis. La información del error se guarda en AnalysisError.txt, que se encuentra en el nivel raíz de la carpeta MiSeqAnalysis.

[Página en blanco intencionada]

Archivos de resultados de análisis del kit universal 1.0

Tipos de archivo de resultados de análisis	46
Formato de archivo BAM	47
Formato de archivo VCF	48
Archivo de cobertura de amplicones	51
Archivos de resultados complementarios	52



Tipos de archivo de resultados de análisis

En la siguiente tabla se describen los archivos de resultados generados para el kit universal 1.0, que proporcionan resultados de análisis para la alineación, las llamadas de variantes y la cobertura.

Nombre de archivo	Descripción
Archivos *.bam	Contienen las lecturas alineadas de una muestra específica. Se encuentra en Data\Intensities\BaseCalls\Alignment.
Archivos *.vcf	Contiene información sobre las variantes que se encuentran en posiciones específicas de un genoma de referencia. Se encuentra en Data\Intensities\BaseCalls\Alignment.
AmpliconCoverage_M#.tsv	Contiene información sobre la cobertura resultante por amplicón por muestra. M# hace referencia al número de manifiesto. Se encuentra en Data\Intensities\BaseCalls\Alignment.



NOTA

El proceso de análisis que genera estos archivos de resultados no es idéntico entre ensayos de FQ y el kit universal 1.0. Esta sección describe únicamente los archivos de resultados para el kit universal 1.0.

Formato de archivo BAM

Un archivo BAM (*.bam) es la versión binaria comprimida de un archivo SAM que se utiliza para representar secuencias alineadas. Los formatos SAM y BAM se describen detalladamente en el sitio web de SAM Tools: samtools.sourceforge.net.

Los archivos BAM se copian en la carpeta de alineación de Data\Intensities\BaseCalls\Alignment con el formato de denominación de archivos SampleName_S#.bam, donde # es el número de la muestra determinado por el orden en que aparecen las muestras en la hoja de muestras.

Los archivos BAM contienen una sección de encabezado y una sección de alineaciones:

- ▶ **Encabezado:** contiene información sobre todo el archivo como, por ejemplo, el nombre de la muestra y la longitud de la muestra. Las alineaciones de la sección de alineaciones están asociadas a la información específica de la sección del encabezado.
- ▶ **Alineaciones:** contiene el nombre de la lectura, la secuencia de lectura, la calidad de la lectura y marcadores personalizados.

Figura 20 Ejemplo de sección de alineación de archivo BAM

```
GA23_40:8:1:10271:11781 64 chr22 17552189 8 35M * 0 0
TACAGACATCCACCACCACACCCAGCTAATTTTTG
IIIII>FA?C::B=:GGGB>GGGEGIIIIHI3EEE#
BC:Z:ATCACG XD:Z:55 SM:I:8
```

El nombre de la lectura incluye el cromosoma y la coordenada de inicio (**chr22 17552189**), la calidad de la alineación (**8**) y el descriptor de la coincidencia (**35M * 0 0**).

Los archivos BAM se pueden visualizar con un visor externo como IGV o UCSC Genome Browser.

Formato de archivo VCF

VCF es un formato de archivo de uso frecuente que ha desarrollado la comunidad de científicos de genómica que contiene información sobre las variantes que se encuentran en posiciones específicas en un genoma de referencia.

Los archivos VCF emplean el formato de denominación de archivos `SampleName_S#.vcf`, donde # es el número de la muestra determinado por el orden en que aparecen las muestras en la hoja de muestras.

- ▶ **Encabezado de archivo VCF:** incluye la versión de formato de archivo VCF y la versión del llamador de variantes. El encabezado enumera las anotaciones utilizadas en el resto del archivo. La última línea del encabezado incluye los encabezados de columna para las líneas de datos. Para obtener más información, consulte *Anotaciones y encabezados de archivos VCF* en la página 49.

Figura 21 Ejemplo de encabezado de archivo VCF

```
##fileformat=VCFv4.1
##FORMAT=<ID=GQX,Number=1,Type=Integer,Description="Minimum of
  {Genotype quality assuming variant position,Genotype quality
  assuming non-variant position}">
##FORMAT=<ID=GT,Number=1,Type=String,Description="Genotype">
##FORMAT=<ID=GQ,Number=1,Type=Float,Description="Genotype
  Quality">
##FORMAT=<ID=AD,Number=.,Type=Integer,Description="Allelic depths
  for the ref and alt alleles in the order listed">
##FORMAT=<ID=VF,Number=1,Type=Float,Description="Variant
  Frequency, the ratio of the sum of the called variant depth to
  the total depth">
##INFO=<ID=TI,Number=.,Type=String,Description="Transcript ID">
##INFO=<ID=GI,Number=.,Type=String,Description="Gene ID">
##INFO=<ID=EXON,Number=0,Type=Flag,Description="Exon Region">
##INFO=<ID=FC,Number=.,Type=String,Description="Functional
  Consequence">
##INFO=<ID=AC,Number=A,Type=Integer,Description="Allele count in
  genotypes, for each ALT allele, in the same order as listed">
##INFO=<ID=AF,Number=A,Type=Float,Description="Allele Frequency,
  for each ALT allele, in the same order as listed">
##INFO=<ID=AN,Number=1,Type=Integer,Description="Total number of
  alleles in called genotypes">
##INFO=<ID=DP,Number=1,Type=Integer,Description="Approximate read
  depth; some reads may have been filtered">
##FILTER=<ID=LowVariantFreq,Description="Low variant frequency <
  0.20">
##FILTER=<ID=LowGQ,Description="GQ below < 20.00">
##FILTER=<ID=LowQual,Description="QUAL below < 100.00">
##FILTER=<ID=R8,Description="IndelRepeatLength is greater than
  8">
##fileDate=20130506
##source=Starling 0.3
##phasing=none
#CHROM POS ID REF ALT QUAL FILTER INFO FORMAT
```

- ▶ **Líneas de datos de archivo VCF:** contiene información sobre una variante individual. Las líneas de datos se incluyen en los encabezados de columna del encabezado.

Anotaciones y encabezados de archivos VCF

El formato de archivo VCF es flexible y ampliable. En las siguientes tablas, se describen las anotaciones y los encabezados de archivos VCF que genera MiSeq Reporter.

Encabezados de archivo VCF

Encabezado	Descripción
CHROM	El cromosoma del genoma de referencia. Los cromosomas aparecen en el mismo orden que el archivo FASTA de referencia.
POS	La posición de base única de la variante en el cromosoma de referencia. En el caso de los SNP, esta posición es la base de referencia con la variante. En el caso de las inserciones y deleciones o las deleciones, esta posición es la base de referencia justo antes de la variante.
ID	El número rs del SNP obtenido a partir del dbSNP.txt, si procede. Si existen varios números rs en esta posición, la lista queda delimitada por punto y coma. Si no existen entradas de dbSNP en esta posición, se utiliza un marcador de valor ausente ('.').
REF	El genotipo de referencia. Por ejemplo, una deleción de una única T se representa como TT de referencia y T alternativa.
ALT	Los alelos que difieren de la lectura de referencia. Por ejemplo, una inserción de una única T se representa como A de referencia y AT alternativa.
QUAL	Una puntuación de calidad según la escala Phred asignada por el llamador de variantes. Las puntuaciones más altas indican una mayor seguridad en la variante y una menor probabilidad de errores. Para una puntuación de calidad de Q, la probabilidad estimada de un error es $10^{-(Q/10)}$. Por ejemplo, el conjunto de llamadas de Q30 tiene un índice de error del 0,1 %. Muchos llamadores de variantes asignan puntuaciones de calidad según sus modelos estadísticos, que son altas en relación con el índice de error observado.

Anotaciones de archivos VCF

Encabezado	Descripción
FILTER (FILTRO)	<p>Si se superan todos los filtros, aparece PASS (APTO) en la columna de filtrado.</p> <ul style="list-style-type: none"> • LowDP (DP baja): se aplica a los lugares con una profundidad de cobertura por debajo del corte. • LowGQ (CG baja): la calidad del genotipado (CG) se encuentra por debajo del corte. • LowQual (CAL baja): la calidad de la variante (CAL) se encuentra por debajo del corte. • LowVariantFreq (Frecuencia de variante baja): la frecuencia de la variante es inferior al umbral. • R8: en el caso de una inserción y deleción, el número de repeticiones adyacentes (una o dos bases) en la referencia es superior a ocho.

Encabezado	Descripción
INFO (INFORMACIÓN)	<ul style="list-style-type: none"> • AC: recuento de alelos en genotipos de cada alelo ALT, en el mismo orden en que se enumeran. • AF: frecuencia de alelos de cada alelo ALT, en el mismo orden en que se enumeran. • AN: número total de alelos en genotipos llamados. • Exon: lista de valores separados por comas que incluye las regiones de exones leídos en RefGene. • FC: consecuencia funcional. • GI: lista de valores separados por comas que incluye los ID de genes leídos en RefGene. • TI: lista de valores separados por comas que incluye los ID de transcripción leídos en RefGene.
FORMAT (FORMATO)	<ul style="list-style-type: none"> • AD: entrada con formato X,Y, donde X es el número de llamadas de referencia e Y es el número de llamadas alternativas. • DP: profundidad de lectura aproximada; se filtran las lecturas con MQ = 255 o con emparejamientos erróneos. • GQ: calidad del genotipo. • GQX: calidad del genotipo. GQX es el mínimo del valor de GQ y la columna QUAL. En general, estos valores son similares. Si se tiene en cuenta el mínimo, GQX indica la medición más fiable relativa a la calidad del genotipo. • GT: genotipo. 0 corresponde a la base de referencia, 1 corresponde a la primera entrada en la columna ALT y, así, sucesivamente. La barra inclinada (/) indica que no hay información de hebras retrasadas disponible. • VF: frecuencia de variante. Se trata del porcentaje de lecturas del alelo alternativo.
SAMPLE (MUESTRA)	La columna SAMPLE (MUESTRA) incluye los valores especificados en la columna FORMAT (FORMATO).

Archivo de cobertura de amplicones

Se genera un archivo de cobertura de amplicones por cada manifiesto. El M# del nombre del archivo representa el número de manifiesto tal y como se muestra en la hoja de muestras.

Cada archivo comienza con una fila de encabezamiento que contiene los ID de las muestras asociados al manifiesto. La primera columna contiene los ID de los objetivos. Cada columna adicional incluye la profundidad de la cobertura del ID de la muestra asociado.

Archivos de resultados complementarios

Los siguientes archivos de resultados proporcionan información complementaria. También pueden incluir resúmenes de errores de análisis y resultados de experimentos. Aunque estos archivos no resultan necesarios para evaluar los resultados de los análisis, se pueden utilizar para la solución de posibles problemas.

Nombre de archivo	Descripción
AnalysisLog.txt	Registro de procesamiento que describe todos los pasos que se han seguido durante el análisis de la carpeta del experimento actual. Este archivo no contiene mensajes de error. Se encuentra en el nivel raíz de la carpeta del experimento.
AnalysisError.txt	Registro de procesamiento que enumera todos los errores que se han producido durante el análisis. Este archivo solo se genera si se producen errores. Se encuentra en el nivel raíz de la carpeta del experimento.
AmpliconRunStatistics.xml	Contiene las estadísticas de resumen específicas del experimento. Se encuentra en el nivel raíz de la carpeta del experimento.
CompletedJobInfo.xml	Archivo que se escribe una vez finalizado el análisis. Contiene información sobre el experimento como, por ejemplo, la fecha, el ID de la celda de flujo, la versión de software y otros parámetros. Se encuentra en el nivel raíz de la carpeta del experimento.
DemultiplexSummaryF1L1.txt	Informa de los resultados del demultiplexado en una tabla con una fila por cada placa y una columna por cada muestra. Se encuentra en Data\Intensities\BaseCalls\Alignment.
ErrorsAndNoCallsByLaneTileReadCycle.csv	Archivo de valores separados por comas que contiene el porcentaje de errores y las ausencias de llamadas de cada placa, lectura y ciclo. Se encuentra en Data\Intensities\BaseCalls\Alignment.
Mismatch.htm	Contiene los histogramas de las incoherencias por ciclo y las ausencias de llamadas por ciclo de cada placa. Se encuentra en Data\Intensities\BaseCalls\Alignment.
Summary.xml	Contiene un resumen de tasas de incoherencia y otros resultados de llamadas de bases. Se encuentra en Data\Intensities\BaseCalls\Alignment.
Summary.htm	Contiene una página web de resumen generada a partir del archivo Summary.xml. Se encuentra en Data\Intensities\BaseCalls\Alignment.

*

*.bam 47
*.bam.bai 47
*.vcf 48

A

alineación 16
análisis
 durante secuenciación 2
AnalysisError.txt 43
AnalysisLog.txt 43
archivo de manifiesto 4, 10
archivos BAM
 formato de archivo 47
archivos de entrada 21
archivos de índice BAM 47
archivos de registro 43
Archivos FASTQ 16
archivos VCF
 anotaciones 49
ArchivosVCF
 formato de archivo 48
asistencia al cliente 55
asistencia técnica 55
ayuda, técnica 55

B

base de datos dbsnp 21
base de datos refGene 21
bases de datos, instaladas
 previamente 21

C

carpeta de análisis 9, 17
carpeta de copia 9
carpeta de datos 9
carpeta de experimento
 acerca de 4
 relación 17
carpeta MiSeqAnalysis 17
carpeta MiSeqOutput 17
ciclos de lectura 9
Cuenta de sistema local 40

D

demultiplexado 16
dirección IP, MiSeq Reporter 3
DLSO 22
documentación 55

E

edición de la hoja de muestras 11
el servicio no se inicia 42
Ensayo de 139 variantes de FQ 20

Ensayo de secuenciación clínica de
 FQ 20
ensayos 20
error al copiar archivos 42

F

flujo de trabajo
 Custom Amplicon 22
 software MiSeqDx 2
Flujo de trabajo Custom Amplicon 22
frecuencia de variante VF 49

G

genomas de referencia, instalados
 previamente 21
genotipo GT 49
gráfico de cobertura 8
gráfico de grupos 7
gráfico de incoherencias 7
gráfico de porcentajes altos 7
gráfico de porcentajes bajos 7
gráfico de puntuación de variantes 8
gráfico Qscore 8
grupos que superan el filtro 14

H

hebras retrasadas, hebras
 adelantadas 14
hoja de muestras
 acerca de 4
 edición 11

I

iconos, estado de análisis 6
ID de gen GI 49
ID de transcripción de TI 49
iniciar sesión como servicio 39
instalación, fuera del instrumento 39

K

kit 20
Kit universal 1.0 20

L

licencia (EULA) 38
localhost 3
LowDP 49
LowGQ 49
LowVariantFreq 49

M

MiSeqDxCf139VariantAssay.txt 35
MiSeqDxCfClinicalSequencingAssay.txt
 35

P

política de seguridad local 39
probabilidad de error 14
Puesta de un análisis en cola por
segunda vez 13
puntuaciones de calidad 14
puntuaciones Q 14

R

r8 49
requisitos informáticos 38
RTAComplete.txt 21
RunInfo.xml 21
ruta del genoma 39
ruta del repositorio 5, 39

S

SAM tools 47
SampleSheet.csv 21
servicio de Windows
acerca de 2
Iniciar sesión como servicio 42
Smith-Waterman 16, 22
solución de problemas
archivos de registro 43
el servicio no se inicia 42
error al copiar archivos 42
superación del filtro (PF) 14

T

tabla de muestras 8
tabla de variantes 8

U

ULSO 22
URL del servidor 5

V

visualización de MiSeq Reporter 3
volver a poner en cola un análisis 6, 11

Asistencia técnica

Si necesita asistencia técnica, póngase en contacto con el servicio de asistencia técnica de Illumina.

Tabla 4 Información de contacto general de Illumina

Sitio web	www.illumina.com
Correo electrónico	techsupport@illumina.com

Tabla 5 Números del servicio de asistencia al cliente de Illumina

Zona	Número de contacto	Zona	Número de contacto
Norteamérica	18008094566	Irlanda	1800812949
Alemania	0800.180.8994	Italia	800874909
Australia	1.800.775.688	Noruega	80016836
Austria	0800296575	Nueva Zelanda	0800451650
Bélgica	080081102	Países Bajos	0800.0223859
Dinamarca	80882346	Reino Unido	0800.917.0041
España	900812168	Suecia	020790181
Finlandia	0800918363	Suiza	0800563118
Francia	0800.911850	Otros países	+44.1799.534000

Hojas de datos de seguridad

Las hojas de datos de seguridad (SDS) están disponibles en el sitio web de Illumina en support.illumina.com/sds.ilmn.

Documentación del producto

En el sitio web de Illumina es posible descargar documentos en PDF sobre el producto. Vaya a support.illumina.com, seleccione un producto y, a continuación, **Documentation & Literature** (Documentación y bibliografía).

