

Questo documento e il suo contenuto sono di proprietà di Illumina, Inc. e delle aziende ad essa affiliate ("Illumina") e sono destinati esclusivamente ad uso contrattuale da parte dei clienti di Illumina, per quanto concerne l'utilizzo dei prodotti qui descritti, con esclusione di qualsiasi altro scopo. Questo documento e il suo contenuto non possono essere usati o distribuiti per altri scopi e/o in altro modo diffusi, resi pubblici o riprodotti, senza previa approvazione scritta da parte di Illumina. Mediante questo documento, Illumina non trasferisce a terzi alcuna licenza ai sensi dei suoi brevetti, marchi, copyright, o diritti riconosciuti dal diritto consuetudinario, né diritti simili di alcun genere.

Al fine di assicurare un uso sicuro e corretto dei prodotti qui descritti, le istruzioni riportate in questo documento devono essere scrupolosamente ed esplicitamente seguite da personale qualificato e adeguatamente addestrato. Leggere e comprendere a fondo tutto il contenuto di questo documento prima di usare tali prodotti.

LA LETTURA INCOMPLETA DEL CONTENUTO DEL PRESENTE DOCUMENTO E IL MANCATO RISPETTO DI TUTTE LE ISTRUZIONI IVI CONTENUTE PUÒ CAUSARE DANNI AL PRODOTTO, LESIONI PERSONALI A UTENTI E TERZI E DANNI MATERIALI.

ILLUMINA NON SI ASSUME ALCUNA RESPONSABILITÀ DERIVANTE DALL'USO IMPROPRIO DEL/DEI PRODOTTO/I QUI DESCRITTI (INCLUSI SOFTWARE O PARTI DI ESSO).

© 2016 Illumina, Inc. Tutti i diritti riservati.

Illumina, MiSeq, MiSeqDx, la tonalità di arancione e la grafica del fluire delle basi sono marchi di fabbrica di Illumina, Inc. e/o delle sue affiliate negli Stati Uniti e/o in altri paesi. Tutti gli altri nomi, loghi e altri marchi di fabbrica sono di proprietà dei rispettivi titolari.

Cronologia revisioni

Documento n.	Data	Descrizione della modifica
Documento n. 15038356 v02	Novembre 2017	Marchi normativi aggiornati.
Documento n. 15038356 v01	Dicembre 2016	Corrette le informazioni nella tabella Targets (Target) per Universal Kit 1.0 modificando inserzioni in delezioni nella riga Deletions (Delezioni). Rimosso Chrome dall'elenco dei browser supportati per l'utilizzo di MiSeq Reporter su un altro computer. Corretti errori di formattazione.
N. codice 15038356 Rev. A	Marzo 2014	Versione iniziale

Sommario

Cronologia revisioni	iii
Sommario	iv
Capitolo 1 Descrizione generale	1
Introduzione	2
Visualizzazione di MiSeq Reporter	3
Concetti di MiSeq Reporter	4
Interfaccia di MiSeq Reporter	5
Rimessa in coda di un'analisi	13
Metriche dell'analisi	14
Procedure di analisi	16
Cartella MiSeqAnalysis	17
Capitolo 2 Visualizzazione dei dati	19
Introduzione	20
Requisiti dei file di input	21
Flusso di lavoro Custom Amplicon (Amplicone personalizzato)	22
File di output dell'analisi per i saggi CF	35
Capitolo 3 Installazione e risoluzione dei problemi	37
Requisiti per l'installazione di MiSeq Reporter su un altro computer	38
Installazione di MiSeq Reporter su un altro computer	39
Utilizzo di MiSeq Reporter su un altro computer	41
Risoluzione dei problemi di MiSeq Reporter	42
Appendice A File di output dell'analisi per Universal Kit 1.0	45
Tipi di file di output dell'analisi	46
Formato file BAM	47
Formato file VCF	48
File di copertura dell'amplicone	51
File di output supplementari	52
Indice	53
Assistenza tecnica	55

[Questa pagina è stata lasciata intenzionalmente vuota]

Descrizione generale

Introduzione	2
Visualizzazione di MiSeq Reporter	3
Concetti di MiSeq Reporter	4
Interfaccia di MiSeq Reporter	5
Rimessa in coda di un'analisi	13
Metriche dell'analisi	14
Procedure di analisi	16
Cartella MiSeqAnalysis	17



Introduzione

Lo strumento MiSeqDx™ è costituito da tre applicazioni software che operano in sequenza per produrre immagini dei cluster sulla cella a flusso, eseguire l'analisi delle immagini e l'identificazione delle basi ed effettuare l'analisi secondaria integrata sullo strumento.

- ▶ Durante la corsa, MiSeq Operating Software (MOS) acquisisce le immagini dei cluster sulla cella a flusso per l'analisi delle immagini, inoltre lavora sul piano portacelle, invia i comandi per dispensare i reagenti e modifica le temperature della cella a flusso.
- ▶ Il software integrato per l'analisi primaria in tempo reale (Real Time Analysis, RTA) esegue l'analisi delle immagini e l'identificazione delle basi e assegna un punteggio qualitativo a ciascuna base per ciascun ciclo man mano che la corsa avanza. Il completamento dell'analisi primaria mediante RTA avvia MiSeq Reporter per iniziare l'analisi secondaria.
- ▶ MiSeq Reporter esegue sullo strumento l'analisi secondaria sulle identificazioni delle basi e sui punteggi qualitativi generati da RTA durante la corsa di sequenziamento. MiSeq Reporter viene eseguito come un servizio Windows e viene visualizzato come un browser web. In alternativa può essere installato su un altro computer. Per ulteriori informazioni, vedere *Installazione di MiSeq Reporter su un altro computer* a pagina 39.

Informazioni sulle applicazioni del servizio Windows

Le applicazioni di servizio Windows service eseguono funzioni specifiche senza l'intervento dell'utente e continuano a funzionare in background durante l'esecuzione di Windows. Poiché opera come applicazione del servizio Windows, MiSeq Reporter avvia automaticamente l'analisi secondaria al termine dell'analisi primaria.

Sequenziamento durante l'analisi

Le risorse di calcolo dello strumento MiSeqDx sono dedicate al sequenziamento oppure all'analisi. Se si avvia un'altra corsa di sequenziamento su MiSeqDx prima che l'analisi secondaria di una corsa precedente sia stata completata, viene visualizzata una finestra di dialogo di conferma. Dopo aver confermato la corsa di sequenziamento, l'analisi secondaria si arresta.

Per riavviare l'analisi secondaria, utilizzare la funzione **Requeue** (Rimetti in coda) sull'interfaccia di MiSeq Reporter una volta completata la nuova corsa di sequenziamento. In tal caso, l'analisi secondaria riprende dall'inizio.

Visualizzazione di MiSeq Reporter

L'interfaccia di MiSeq Reporter può essere visualizzata solo mediante un browser Web. Per visualizzare l'interfaccia di MiSeq Reporter, aprire un qualsiasi browser Web su un computer con accesso alla stessa rete dello strumento MiSeqDx. Quindi collegarsi al servizio HTTP sulla porta **8042** mediante uno dei metodi seguenti:

- ▶ Connessione utilizzando l'indirizzo IP dello strumento seguito da 8042.

Indirizzo IP	Porta del servizio HTTP	Indirizzo HTTP
10.10.10.10, ad esempio	8042	10.10.10.10:8042

- ▶ Connessione utilizzando un nome di rete per MiSeqDx seguito da 8042

Nome della rete	Porta del servizio HTTP	Indirizzo HTTP
MiSeqDx01, ad esempio	8042	MiSeqDx01:8042

Per le installazioni di MiSeq Reporter su un altro computer, collegarsi con il metodo utilizzato per le applicazioni di servizio installate sullo strumento, **localhost** seguito da 8042.

Su un altro computer	Porta del servizio HTTP	Indirizzo HTTP
localhost	8042	localhost:8042

Per ulteriori informazioni, vedere *Installazione di MiSeq Reporter su un altro computer* a pagina 39.

Concetti di MiSeq Reporter

I seguenti concetti e i termini sono utilizzati in MiSeq Reporter.

Concetto	Descrizione
Manifest (File manifest)	Il file che specifica un genoma di riferimento e le regioni di riferimento target da usare nella fase di allineamento. Il file manifest utilizzato dai saggi per fibrosi cistica è pre-caricato su MiSeqDx.
Repository	Cartella contenente i dati generati durante le corse di sequenziamento. Ciascuna cartella delle corse è una sottocartella del repository.
Cartella delle corse	La struttura della cartella popolata dal software di analisi primaria RTA (cartella MiSeqOutput) o la cartella popolata da MiSeq Reporter (MiSeqAnalysis).
Foglio campioni	Un file con valori separati da virgola (comma-separated values, *.csv) contenente le informazioni necessarie per impostare e analizzare una corsa di sequenziamento, comprendente anche un elenco di campioni e le relative sequenze indice. Il file viene creato su un altro computer mediante il software Illumina Worklist Manager. Il foglio campioni deve essere fornito durante le fasi d'impostazione della corsa su MiSeqDx. Dopo l'inizio della corsa, il foglio campioni viene automaticamente rinominato SampleSheet.csv e copiato nelle cartelle delle corse: MiSeqOutput and MiSeqAnalysis.
Flusso di lavoro	Una procedura di analisi secondaria eseguita da MiSeq Reporter. Il flusso di lavoro per ogni corsa è specificato nel foglio campioni.

Interfaccia di MiSeq Reporter

Quando MiSeq Reporter si apre nel browser, viene visualizzata la schermata principale con un'immagine dello strumento al centro. Le icone Settings (Impostazioni) e Help (Guida) si trovano nell'angolo in alto a destra, mentre nell'angolo in alto a sinistra si trova la scheda Analyses (Analisi).



- ▶ **MiSeq Reporter Help** (Guida di MiSeq Reporter): selezionare l'icona Help (Guida) per aprire la documentazione di MiSeq Reporter nella finestra del browser.
- ▶ **Settings** (Impostazioni): selezionare l'icona Settings (Impostazioni)  per modificare il percorso dell'URL del server e del repository.
- ▶ Scheda **Analyses** (Analisi): selezionare Analyses (Analisi) per espandere la scheda. Nella scheda Analyses (Analisi) è riportato un elenco di corse di analisi completate, messe in coda o in corso.

Figura 1 Schermata principale di MiSeq Reporter

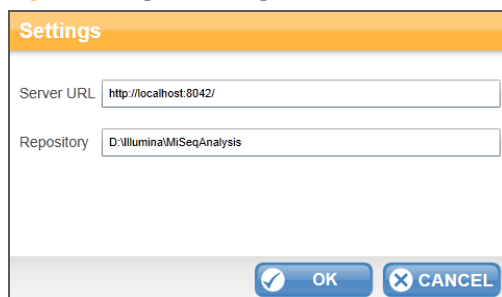


Impostazioni dell'URL del server o del repository

Per modificare l'URL del server e il percorso del repository, utilizzare la funzione Settings (Impostazioni) :

- ▶ **Server URL** (URL del server): il server su cui viene eseguito MiSeq Reporter.
- ▶ **Repository path** (Percorso del repository): posizione della cartella di analisi in cui vengono scritti i file di output.

Figura 2 Impostazioni per l'URL del server e del repository



Di solito non è necessario modificare queste impostazioni a meno che MiSeq Reporter non venga eseguito su un altro computer.

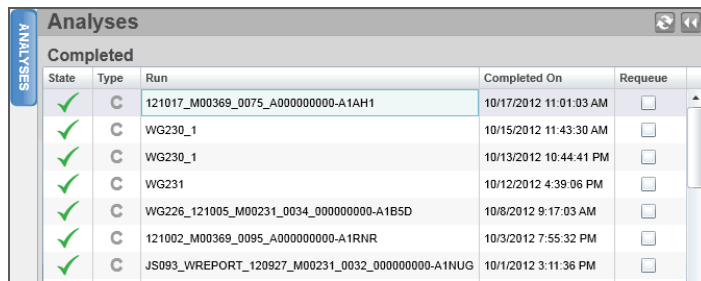
In tal caso, impostare il percorso del repository nella posizione in rete della cartella MiSeqOutput. Per maggiori informazioni, vedere *Utilizzo di MiSeq Reporter su un altro computer* a pagina 41.

Scheda Analyses (Analisi)

Nella scheda Analyses (Analisi) sono elencate tutte le corse di sequenziamento che si trovano nel repository specificato. Da questa scheda è possibile aprire i risultati di qualsiasi corsa fra quelle elencate, oppure è possibile rimettere in coda per l'analisi una corsa selezionata.

Selezionare l'icona **Refresh Analysis List** (Aggiorna elenco analisi)  nell'angolo in alto a destra per aggiornare l'elenco in qualsiasi momento.

Figura 3 Scheda Analyses (Analisi) allargata






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-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"/>
✓	C	JS093_WREPORT_120927_M00231_0032_000000000-A1NUG	10/1/2012 3:11:36 PM	<input type="checkbox"/>

La scheda Analysis (Analisi) è formata dalle colonne State (Stato), Type (Tipo), Run (Corsa), Completed On (Completata il) e Requeue (Rimetti in coda):

- ▶ **State** (Stato): indica la situazione attuale dell'analisi mediante una di tre icone di stato.

Tabella 1 Icone dello stato dell'analisi

Icona	Descrizione
	Indica che l'analisi secondaria è stata completata correttamente.
	Indica che l'analisi secondaria è in corso.
	Indica che si sono verificati errori e l'analisi secondaria non è stata completata correttamente.

- ▶ **Type** (Tipo): elenca il flusso di lavoro dell'analisi associato a ogni corsa mediante un indicatore rappresentato da una sola lettera. Per i saggi della fibrosi cistica e Universal Kit 1.0, l'indicatore è la lettera C.
- ▶ **Run** (Corsa): il nome della cartella delle corse nelle cartelle MiSeqOutput e MiSeqAnalysis.
- ▶ **Completed On** (Completata il): la data in cui è stata completata l'analisi secondaria.
- ▶ **Requeue** (Rimetti in coda): selezionare la casella di controllo per rimettere in coda una particolare procedura da analizzare. Viene visualizzato il pulsante Requeue (Rimetti in coda). Per maggiori informazioni, vedere *Rimessa in coda di un'analisi* a pagina 13.





Quando si rimette in coda un'analisi, la corsa viene visualizzata nella parte inferiore della scheda Analyses (Analisi) e l'icona  indica che è in corso.

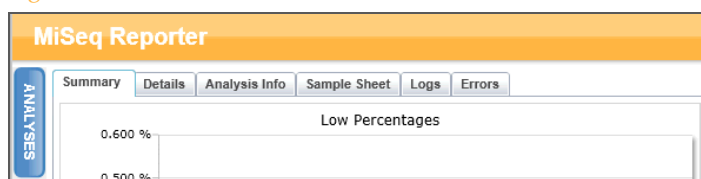
Figura 4 Corsa in coda nella scheda Analyses (Analisi)

	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	

Schede di informazioni e risultati delle analisi

Dopo aver selezionato una corsa dalla scheda Analyses (Analisi), le informazioni e i risultati che la riguardano vengono visualizzati in una serie di schede sull'interfaccia MiSeq Reporter: Summary, Details, Analysis Info, Sample Sheet, Logs e Errors (Riepilogo, Dettagli, Informazioni sull'analisi, Foglio campioni, Registri e Errori). Inizialmente vengono visualizzate le informazioni sulle schede Analysis Info (Informazioni sull'analisi) e Sample Sheet (Foglio campioni). Una volta completata l'analisi, tutte le schede si riempiono.

Figura 5 Schede di informazioni e risultati

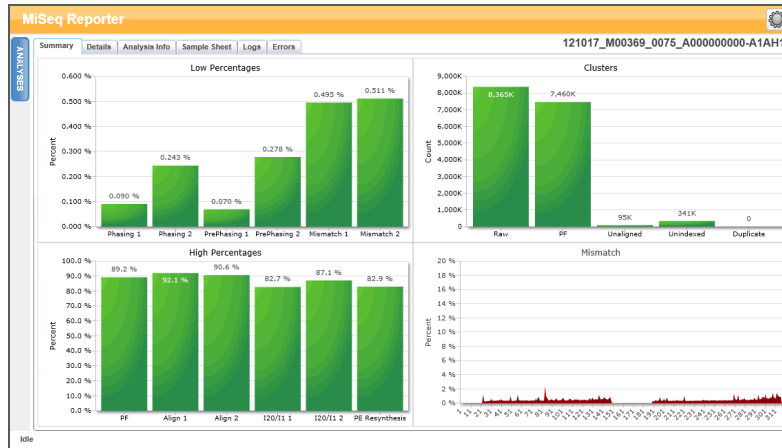


Scheda Summary (Riepilogo)

Nella scheda Summary (Riepilogo) sono riepilogati i risultati dell'analisi. Nella scheda Summary (Riepilogo) sono rappresentati quattro grafici:

- ▶ Grafico **Low Percentages** (Percentuali basse): mostra la determinazione delle fasi (phasing), la predeterminazione delle fasi (prephasing) e le mancate corrispondenze espresse in percentuali. Percentuali basse indicano che le statistiche della corsa sono buone. Per maggiori informazioni, vedere *Determinazione delle fasi (phasing) e predeterminazione delle fasi (prephasing)* a pagina 15.
- ▶ Grafico **High Percentages** (Percentuali alte): mostra i cluster che attraversano il filtro, l'allineamento a un riferimento e le intensità espressi in percentuali. Percentuali alte indicano che le statistiche della corsa sono buone.
- ▶ Grafico **Clusters** (Cluster): riporta il numero dei cluster non elaborati, dei cluster che attraversano il filtro, dei cluster che non si allineano, dei cluster non associati a un indice e dei duplicati.
- ▶ Grafico **Mismatch** (Mancata corrispondenza): riporta le mancate corrispondenze per ciascun ciclo. Per mancata corrispondenza si intende una differenza fra la lettura del sequenziamento e un genoma di riferimento dopo l'allineamento.

Figura 6 Scheda Summary (Riepilogo)



Scheda Details (Dettagli)

Nella scheda Details (Dettagli) sono riportati i dettagli dei risultati dell'analisi. Le tabelle e i grafici seguenti possono essere visualizzati nella scheda Details (Dettagli) in base al saggio o al kit usato:

- ▶ Tabella **Samples** (Campioni): riepiloga i risultati del sequenziamento relativi a ciascun campione.
- ▶ Tabella **Targets** (Target): mostra le statistiche per le regioni target di un campione selezionato. Solo per Universal Kit 1.0.
- ▶ Tabella **Variants** (Varianti): mostra le differenze fra il campione di DNA e il riferimento.
- ▶ Grafico **Coverage** (Copertura): mostra la profondità di sequenziamento del campione misurando il numero di basi presenti nella sequenza campione per ciascuna posizione del riferimento.
- ▶ Grafico **Qscore** (Punteggio qualitativo): mostra il punteggio qualitativo medio, vale a dire la probabilità stimata di errore di un'identificazione delle basi. Per maggiori informazioni vedere *Grafico Qscore (Punteggio qualitativo)* a pagina 33.
- ▶ Grafico **Variant Score** (Punteggio varianti): mostra la posizione delle SNV e delle Indel.

Figura 7 Esempio di scheda Details (Dettagli) per il saggio CF 139-Variant

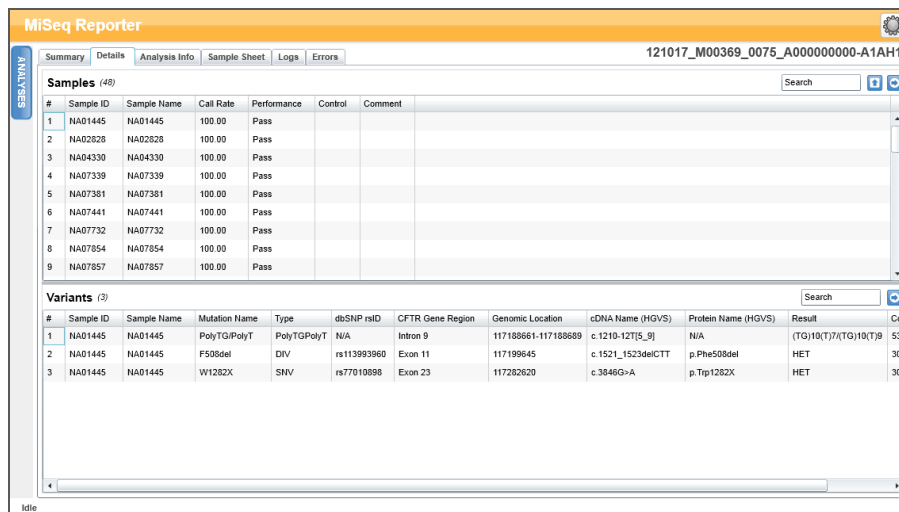
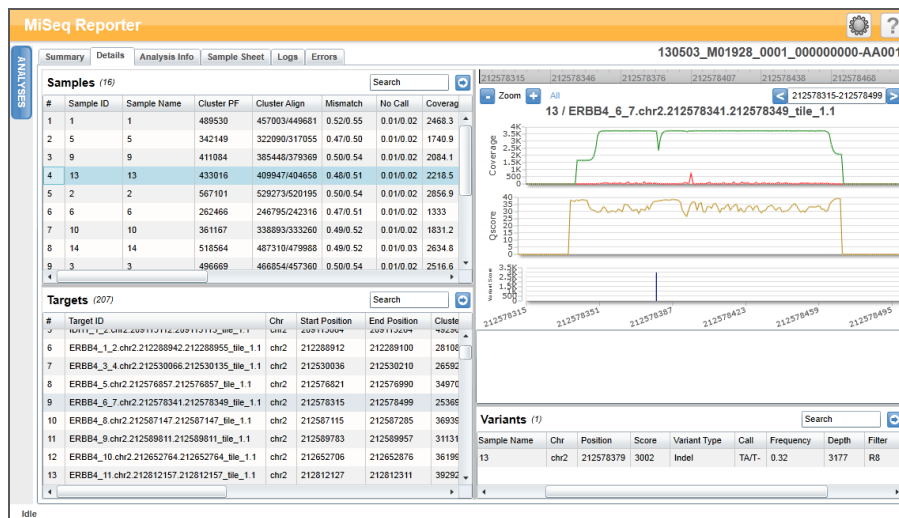


Figura 8 Esempio di scheda Details (Dettagli) per Universal Kit 1.0



- 

I risultati delle tabelle Samples (Campioni), Targets (Target) o Variants (Varianti) possono essere esportati singolarmente in un file di testo mediante l'icona **Export table data to text file** (Esporta dati della tabella in un file di testo). L'esportazione non altera il file report dell'analisi.
- 

Per i saggi CF, i risultati possono essere esportati in un file report dell'analisi di CF mediante l'icona **Export data to CF report** (Esporta dati in report CF). Per maggiori informazioni, vedere *File di output dell'analisi per i saggi CF* a pagina 35.

Scheda Analysis Info (Informazioni sull'analisi)

La scheda Analysis Info (Informazioni sull'analisi) contiene informazioni logistiche relative alla corsa e all'analisi.

Figura 9 Scheda Analysis Info (Informazioni sull'analisi)

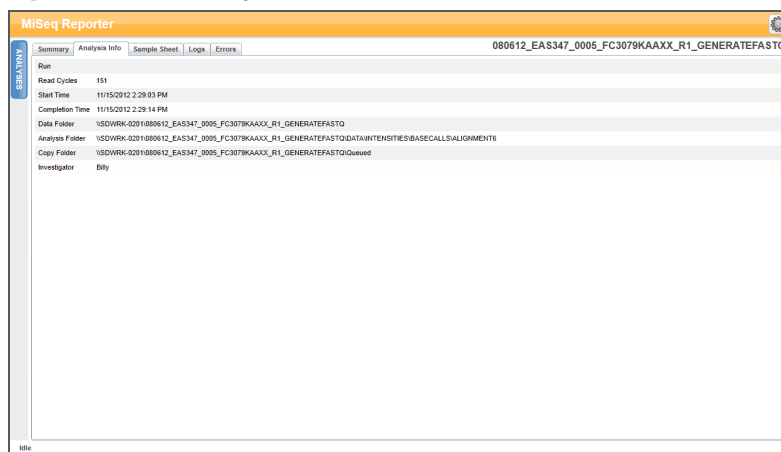


Tabella 2 Scheda Analysis Info (Informazioni sull'analisi)

Riga	Descrizione
Read Cycles (Cicli di lettura)	Rappresentazione del numero di cicli in ogni lettura, compresa l'annotazione per le letture indici. Ad esempio, una corsa annotata come 151, 8 (I), 8 (I), 151 indica 151 cicli per la prima lettura, 8 cicli per la prima lettura indici, 8 cicli per la seconda lettura indici e una lettura finale di 151 cicli.
Start Time (Ora d'inizio)	Ora in cui è stata avviata l'analisi secondaria.
Completion Time (Ora di completamento)	Ora in cui è stata completata l'analisi secondaria.
Data Folder (Cartella dati)	Livello della radice della cartella di output prodotta dal software di analisi primaria RTA (MiSeqOutput) che contiene tutti gli output di analisi primaria e secondaria relativi alla corsa.
Analysis Folder (Cartella analisi)	Percorso completo per la cartella Alignment (Allineamento) nella cartella MiSeqAnalysis (Data \ Intensities \ BaseCalls \ Alignment).
Copy Folder (Cartella copia)	Percorso completo della sottocartella Queued (In coda) nella cartella MiSeqAnalysis.

Scheda Sample Sheet (Foglio campioni)

Nella scheda Sample Sheet (Foglio campioni) sono riportati i parametri della corsa specificati nel foglio campioni oltre agli strumenti per modificare il foglio e quindi rimettere in coda la corsa.

Figura 10 Esempio di scheda Sample Sheet (Foglio campioni) per Universal Kit 1.0

Sample_ID	Sample_Name	Sample_Well	Control	I7_Index_ID	index	IS_Index_ID	index2	Manifest	GenomeFolder
1		B01		A701	ATCACGAC	A502	TGCTAAGT	A	Home_sapiens\UCSC\hg19\Sequence\WholeGenomeF
5		B02		A702	ACAGTGGT	A502	TGCTAAGT	A	Home_sapiens\UCSC\hg19\Sequence\WholeGenomeF

Tabella 3 Contenuto della scheda Sample Sheet (Foglio campioni)

Riga	Descrizione
Date (Data)	Data in cui è stata eseguita la corsa di sequenziamento.
Workflow (Flusso di lavoro)	Flusso di lavoro dell'analisi per la corsa. Per i saggi CF e Universal Kit 1.0, il nome del flusso di lavoro è Custom Amplicon (Amplicone personalizzato).
Application (Applicazione)	Nome dell'applicazione. Utilizzato da Illumina Worklist Manager, questo campo indica quale saggio o kit è utilizzato per la corsa.
Assay (Saggio)	Nome del saggio o del kit.
Chemistry (Chimica)	Il nome della chimica identifica i frammenti usati per comporre la ricetta specifica della corsa. Per le corse MiSeqDx, il nome della chimica è amplicon (amplicone).
Manifests (File manifest)	Il nome del file manifest che specifica un genoma di riferimento e le regioni di riferimento mirate da usare nella fase di allineamento.
Reads (Lecture)	Numero di cicli eseguiti in Read 1 (Lettura 1) e in Read 2 (Lettura 2). In questa sezione non sono comprese le letture degli indici.
Settings (Impostazioni)	Parametri opzionali della corsa.
Data (Dati)	ID campione, nome del campione, sequenze degli indici e percorso della cartella dei genomi. I requisiti variano a seconda del flusso di lavoro.

Scheda Logs (Registri)

Nella scheda Logs (Registri) sono elencati tutti i passaggi effettuati durante l'analisi. Tali passaggi vengono registrati nei file log che si trovano nella cartella Logs (Registri) e riepilogati nel file AnalysisLog.txt, un file importante per la risoluzione dei problemi.

Scheda Errors (Errori)

Nella scheda Errors (Errori) sono elencati gli errori verificatisi durante l'analisi e riepilogati nel file AnalysisError.txt, un file importante per la risoluzione dei problemi.

Modifica del foglio campioni in MiSeq Reporter

È possibile modificare i dati del foglio campioni per una specifica corsa dalla scheda Sample Sheet (Foglio campioni) sull'interfaccia web di MiSeq Reporter. Per modificare un foglio campioni sono necessari un mouse e una tastiera.



ATTENZIONE

Le modifiche delle informazioni nel foglio campioni devono essere eseguite con la massima cautela e attenzione. Il monitoraggio dei campioni potrebbe essere alterato e condurre eventualmente a report con risultati errati.

- ▶ Per modificare una riga nel foglio campioni, fare clic sul campo e apportare le modifiche necessarie.
- ▶ Per aggiungere una riga al foglio campioni, fare clic sulla riga e selezionare **Add Row** (Aggiungi riga). La nuova riga viene visualizzata sotto la riga selezionata.

ADD ROW

- ▶ Per eliminare una riga dal foglio campioni, fare clic sulla riga e selezionare **Delete Row** (Elimina riga).

A blue rectangular button with rounded corners containing the text "DELETE ROW" in white, uppercase letters.

- ▶ Una volta completate le modifiche nel foglio campioni, selezionare **Save and Requeue** (Salva e rimetti in coda). In questo modo le modifiche saranno salvate e avrà inizio l'analisi secondaria con il foglio campioni modificato.

A blue rectangular button with rounded corners containing the text "SAVE AND REQUEUE" in white, uppercase letters.

- ▶ Se è stata fatta una modifica per errore al foglio campioni, prima di salvare le modifiche, fare clic su una scheda adiacente. Verrà visualizzato un messaggio che avvisa che le modifiche non sono state salvate. Fare clic su **Discard** (Annulla) per annullare le modifiche.

A blue rectangular button with rounded corners containing a white "X" icon followed by the text "DISCARD" in white, uppercase letters.

Salvataggio di grafici come immagini

In MiSeq Reporter è presente l'opzione di salvare un'immagine dei grafici generati per una corsa. Fare clic con il tasto destro del mouse su una qualsiasi posizione della scheda Summary (Riepilogo) o sulla posizione del grafico sulla scheda Details (Dettagli), quindi fare clic con il tasto sinistro su **Save Image As** (Salva immagine con nome). Quando viene richiesto, assegnare un nome al file e selezionare la posizione in cui salvarlo.

Tutte le immagini vengono salvate in formato JPG. Tutti i grafici illustrati nella scheda vengono esportati in un unico grafico. Per utilizzare questa opzione è necessario disporre di un mouse.

Rimessa in coda di un'analisi

È possibile rimettere in coda un'analisi dall'interfaccia Web di MiSeq Reporter. Prima di procedere, verificare che non si stia svolgendo una corsa di sequenziamento.

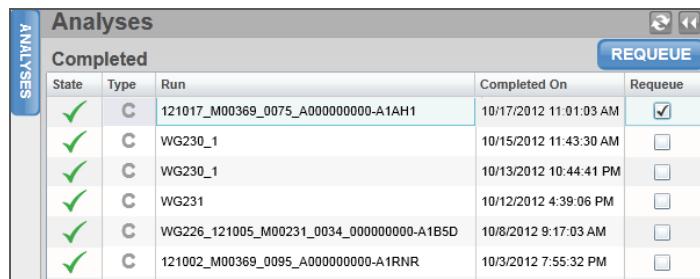
Ogniqualvolta un'analisi viene rimessa in coda, nella cartella MiSeqAnalysis si crea una nuova cartella Alignment (Allineamento) con un numero in sequenza aggiunto al nome della cartella. Per esempio, Alignment (Allineamento), Alignment1 (Allineamento1), Alignment2 (Allineamento2).

MiSeqAnalysis \<RunFolderName> \Data \Intensities \BaseCalls \Alignment2

- 1 Dall'interfaccia Web di MiSeq Reporter, fare clic su **Analyses** (Analisi).
- 2 Individuare la corsa nell'elenco delle corse disponibili e fare clic sulla casella di controllo Requeue (Rimetti in coda) accanto al nome della corsa.

Se la corsa non è nell'elenco, cambiare il repository specificato nella posizione corretta. Per maggiori informazioni, vedere *Impostazioni dell'URL del server o del repository* a pagina 5.

Figura 11 Rimessa in coda di un'analisi



Analyses				
Completed				
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 Fare clic su **Requeue** (Rimetti in coda). L'icona State (Stato) a sinistra del nome della corsa cambia, mostrando che l'analisi è in corso 🔄.



NOTA

Se l'analisi non si avvia, assicurarsi che i seguenti file di input siano presenti nella cartella delle corse dell'analisi: SampleSheet.csv, RTACComplete.txt e RunInfo.xml.

Metriche dell'analisi

Durante la corsa di sequenziamento, l'analisi in tempo reale (RTA) genera file di dati che comprendono le metriche dell'analisi utilizzate da MiSeq Reporter per l'analisi secondaria. Le metriche visualizzate nei report dell'analisi secondaria riguardano i cluster che attraversano il filtro, i punteggi qualitativi di identificazione delle basi e i valori di determinazione delle fasi (phasing) e predeterminazione delle fasi (prephasing).

Cluster che attraversano il filtro

I cluster che attraversano il filtro consentono di misurare la qualità dei cluster. Il filtro rimuove i dati meno affidabili filtrando i dati non elaborati per rimuovere qualsiasi lettura che non soddisfi la qualità complessiva. Nei report dell'analisi, i cluster che attraversano il filtro sono indicati con PF.

Punteggi qualitativi

Un punteggio qualitativo (Q-score) è una previsione della probabilità di un'identificazione delle basi errata. Durante la corsa di sequenziamento, i punteggi qualitativi di identificazione delle basi vengono registrati per ciascun ciclo. Nel corso dell'analisi, i punteggi qualitativi vengono registrati in file FASTQ in formato codice ASCII.

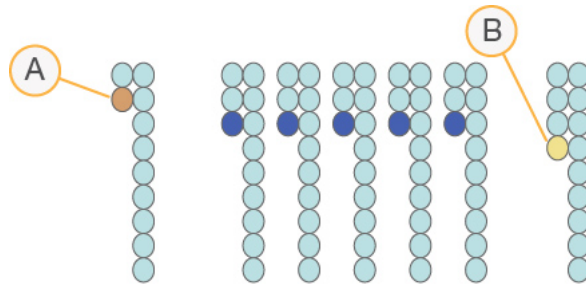
La tabella seguente illustra la relazione fra il punteggio qualitativo e la probabilità di errore.

Punteggio qualitativo	Probabilità di errore
Q40	0,0001 (1 su 10.000)
Q30	0,001 (1 su 1.000)
Q20	0,01 (1 su 100)
Q10	0,1 (1 su 10)

Determinazione delle fasi (phasing) e predeterminazione delle fasi (prephasing)

Durante la reazione di sequenziamento, ciascun filamento di DNA in un cluster si estende di una base per ciclo. Una piccola parte di filamenti potrebbe andare fuori fase con il ciclo di incorporazione corrente, o restando indietro di una base (phasing) o saltando di una base in avanti (prephasing). Le percentuali di determinazione delle fasi (phasing) e predeterminazione delle fasi (prephasing) rappresentano una stima della frazione di molecole che rientrano nella determinazione delle fasi (phasing) e nella predeterminazione delle fasi (prephasing) in ciascun ciclo.

Figura 12 Determinazione delle fasi (phasing) e predeterminazione delle fasi (prephasing)



- A Lettura con una base nella determinazione delle fasi (phasing)
- B Lettura con una base nella predeterminazione delle fasi (prephasing)

Il numero di cicli eseguiti in una lettura è pari a un ciclo in più rispetto al numero di cicli analizzati. Ad esempio, una corsa paired-end da 150 cicli esegue due letture da 151 cicli (2×151) con un totale di 302 cicli. Al termine della corsa, si analizzano 2×150 cicli. Per i calcoli di predeterminazione delle fasi (prephasing) è necessario un ulteriore ciclo per Read 1 (Lettura 1) e Read 2 (Lettura 2).

Procedure di analisi

MiSeq Reporter esegue l'analisi secondaria mediante una serie di procedure di analisi che comprendono il demultiplex, la generazione di file FASTQ, l'allineamento e l'identificazione delle varianti.

Demultiplex

Il demultiplex costituisce la prima fase dell'analisi se nel foglio campioni sono elencati più campioni e la corsa presenta letture indici.

Mediante il demultiplex è possibile separare i dati da campioni raggruppati in base a sequenze indice brevi che etichettano i campioni di diverse librerie. Ciascuna sequenza di lettura indici viene messa a confronto con le sequenze indice specificate nel foglio campioni. In questa fase non vengono considerati i valori qualitativi.

Generazione di file FASTQ

Dopo il demultiplex, questa procedura genera file intermedi in formato FASTQ, un formato di testo utilizzato per rappresentare le sequenze. I file FASTQ rappresentano gli input principali per la fase di allineamento. I file FASTQ contengono le letture di ciascun campione e i punteggi qualitativi, fatta eccezione per le letture dei cluster che non hanno attraversato il filtro.

Allineamento

Mediante l'allineamento è possibile confrontare le sequenze rispetto al riferimento al fine di identificare una relazione fra le sequenze e assegnare un punteggio in base a regioni di similarità. Le letture allineate vengono scritte su file in formato BAM.

Per i dati generati su MiSeq Reporter, MiSeqDx utilizza un algoritmo di Smith-Waterman con matrice a banda che esegue gli allineamenti locali di sequenze per determinare le regioni simili fra due sequenze. Anziché considerare la sequenza intera, l'algoritmo di Smith-Waterman confronta i segmenti di tutte le lunghezze possibili. Gli allineamenti locali sono utili per le sequenze dissimili sospettate di contenere regioni di similarità nella sequenza più ampia.

Identificazione delle varianti

L'identificazione delle varianti registra i polimorfismi di singoli nucleotidi (Single Nucleotide Polymorphism, SNP), inserzioni e delezioni (Indel) e altre varianti strutturali.

Per i dati generati sullo strumento MiSeqDx, l'identificazione delle varianti viene eseguita da Starling Variant Caller in MiSeq Reporter. Starling identifica SNP e Indel piccoli e riepiloga la profondità e le probabilità di errore per ciascun sito nel genoma. Per ciascun SNP o Indel identificato, la probabilità di errore è fornita come un punteggio qualitativo della variante.

Al completamento, Starling produce report di SNP e Indel in formato html e file di testo delimitato da tabulazione contenenti le varianti nel formato di identificazione delle varianti (Variant Call Format, VCF). Per maggiori informazioni, vedere *Formato file VCF* a pagina 48.

Cartella MiSeqAnalysis

La cartella MiSeqAnalysis è la cartella principale delle corse per MiSeq Reporter. Il rapporto esistente fra le cartelle delle corse MiSeqOutput e MiSeqAnalysis si riassume nei seguenti punti:

- ▶ Durante il sequenziamento, l'analisi in tempo reale (RTA) riempie la cartella MiSeqOutput con i file generati durante l'analisi primaria.
- ▶ A eccezione delle immagini di focalizzazione e delle immagini in miniatura, l'RTA copia i file nella cartella MiSeqAnalysis in tempo reale. Una volta completata l'analisi primaria, l'RTA scrive il file RTAComplete.xml in entrambe le cartelle delle corse.
- ▶ MiSeq Reporter controlla la cartella MiSeqAnalysis e avvia l'analisi secondaria quando viene visualizzato il file RTAComplete.xml .
- ▶ Mentre l'analisi secondaria avanza, MiSeq Reporter scrive i file di output dell'analisi nella cartella MiSeqAnalysis, quindi copia i file nella cartella MiSeqOutput.

[Questa pagina è stata lasciata intenzionalmente vuota]

Visualizzazione dei dati

Introduzione	20
Requisiti dei file di input	21
Flusso di lavoro Custom Amplicon (Amplicone personalizzato)	22
File di output dell'analisi per i saggi CF	35



Introduzione

MiSeq Reporter esegue l'analisi secondaria e genera diversi tipi di informazioni specifiche per il saggio al completamento dell'analisi. I risultati vengono visualizzati nell'interfaccia Web di MiSeq Reporter in forma di grafici e tabelle per ciascuna corsa. I prodotti di MiSeqDx sono elencati nella seguente tabella:

Prodotto	Descrizione
Saggio Cystic Fibrosis 139-Variant	Rileva 139 varianti rilevanti dal punto di vista clinico nel gene CFTR da un massimo di 48 campioni.
Saggio Cystic Fibrosis Clinical Sequencing	Rileva le mutazioni nelle regioni che codificano le proteine compresi i confini introne/esone, due ampie delezioni e due mutazioni introniche profonde nel gene CFTR da un massimo di 8 campioni.
Universal Kit 1.0	Set di reagenti e materiali di consumo usati assieme agli oligonucleotidi personalizzati forniti dall'utente per eseguire il risequenziamento mirato di regioni genomiche di interesse specifiche.

Requisiti dei file di input

Per poter eseguire l'analisi secondaria o l'analisi messa in coda, è necessario che MiSeq Reporter disponga dei seguenti file di analisi primaria generati durante la corsa di sequenziamento. Per eseguire l'analisi sono necessari file dell'analisi primaria come *.bcl, *.filter e *.locs.

Non è necessario spostare o copiare i file in un altro percorso prima che l'analisi venga avviata. I file necessari vengono copiati automaticamente nella cartella MiSeqAnalysis durante il sequenziamento.

Nome file	Descrizione
RTAComplete.txt	File di marker che indica che l'elaborazione RTA è stata completata. La presenza di questo file innesca la messa in coda dell'analisi da parte di MiSeq Reporter.
SampleSheet.csv	Fornisce i parametri per la corsa e la successiva analisi. All'avvio della corsa, il foglio campioni viene copiato a livello della radice della cartella della corsa e rinominato come SampleSheet.csv.
RunInfo.xml	Contiene informazioni ad alto livello sulla corsa, ad esempio il numero di letture e cicli nella corsa di sequenziamento e se si tratta o meno di una lettura indicizzata.

Database e genomi preinstallati

MiSeqDx include database e genomi preinstallati.

Preinstallati	Descrizione
Database	dbSNP (database di SNP umani), versione 131 refGene (database di sequenze di riferimento umane)
Genomi	Umano (<i>Homo sapiens</i>) build hg19

Flusso di lavoro Custom Amplicon (Amplicone personalizzato)

Il flusso di lavoro Custom Amplicon (Amplicone personalizzato), usato per i saggi della fibrosi cistica e Universal Kit 1.0, valuta le regioni brevi di DNA amplificato, o ampliconi, per identificare le varianti. Il sequenziamento mirato degli ampliconi consente di coprire un numero elevato di regioni particolari attraverso un numero di campioni.

Dopo il demultiplex e la generazione di file FASTQ, il flusso di lavoro esegue i seguenti passaggi:

- ▶ **Allineamento:** i cluster di ciascun campione vengono allineati rispetto alle sequenze di ampliconi specificate nel file manifest.
 - Per i dati paired-end, ciascuna lettura viene valutata inizialmente in termini di allineamento con le sequenze sonda pertinenti per quella lettura. Read 1 (Lettura 1) è valutata rispetto al complemento inverso degli oligonucleotidi locus specifici a valle (Downstream Locus Specific Oligo, DLSO), mentre Read 2 (Lettura 2) è valutata rispetto agli oligonucleotidi locus specifici a monte (Upstream Locus Specific Oligo, ULSO). Se l'inizio di una sequenza di lettura corrisponde a una sequenza sonda con non più di una non corrispondenza, tutta la lunghezza della lettura viene allineata rispetto alla sequenza target di ampliconi per quella sequenza sonda. Tale allineamento viene eseguito per la lunghezza delle sequenze target di ampliconi mediante un allineamento di Smith-Waterman con matrice a banda.
 - Data la chimica del saggio, non si osservano Indel né nei DLSO né negli ULSO.
- ▶ **Valutazione paired-end:** per le corse paired-end, per ogni lettura si considera l'allineamento con il punteggio massimo. Se una delle letture non è allineata o è allineata a cromosomi diversi, le letture vengono contrassegnate come coppia non risolta. Inoltre, se due allineamenti provengono da ampliconi diversi (vale a dire righe diverse nella sezione target del file manifest), le letture vengono contrassegnate come coppia non risolta.
- ▶ **Raggruppamento/ordinamento:** le letture vengono raggruppate per campione e cromosoma, quindi ordinate in base alla posizione del cromosoma. I risultati vengono scritti su un file BAM per ciascun campione.
- ▶ **Identificazione delle varianti:** le mutazioni vengono identificate mediante Variant Caller. Per maggiori informazioni, vedere *Identificazione delle varianti* a pagina 16.
- ▶ **Analisi e annotazione delle varianti:** grazie a un database SNP (dbsnp.txt) preinstallato, qualsiasi mutazione nota viene contrassegnata nel file report dell'analisi.
- ▶ **Report delle statistiche:** le statistiche vengono riepilogate in un report.

Scheda Summary (Riepilogo)

Le informazioni visualizzate sulla scheda Summary (Riepilogo) comprendono un grafico delle percentuali basse, un grafico delle percentuali alte, un grafico dei cluster e un grafico delle mancate corrispondenze.

Figura 13 Esempio di scheda Summary (Riepilogo)

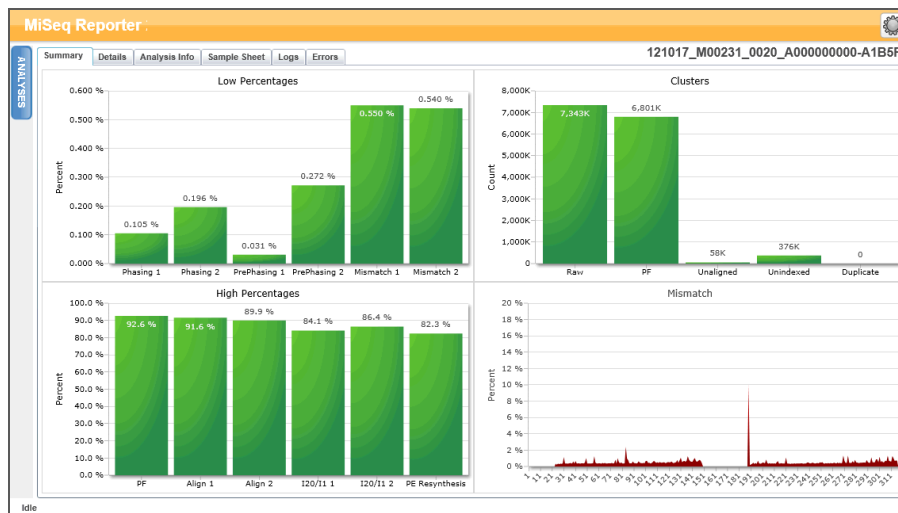


Grafico Low Percentages (Percentuali basse)

Asse Y	Asse X	Descrizione
Percent (Percentuale)	Phasing 1 (Determinazione delle fasi [phasing] 1)	Percentuale di molecole in un cluster oltre l'attuale ciclo in Read 1 (Lettura 1).
	Phasing 2 (Determinazione delle fasi [phasing] 2)	Percentuale di molecole in un cluster oltre l'attuale ciclo in Read 2 (Lettura 2).
	PrePhasing 1 (Predeterminazione delle fasi [prephasing] 1)	Percentuale di molecole in un cluster in avanti rispetto all'attuale ciclo in Read 1 (Lettura 1).
	PrePhasing 2 (Predeterminazione delle fasi [prephasing] 2)	Percentuale di molecole in un cluster in avanti rispetto all'attuale ciclo in Read 2 (Lettura 2).
	Mismatch 1 (Mancata corrispondenza 1)	Percentuale media di mancate corrispondenze per Read 1 (Lettura 1) per tutti i cicli.
	Mismatch 2 (Mancata corrispondenza 2)	Percentuale media di mancate corrispondenze per Read 2 (Lettura 2) per tutti i cicli.

Grafico High Percentages (Percentuali alte)

Asse Y	Asse X	Descrizione
Percent (Percentuale)	PF (Che attraversano il filtro)	Percentuale di cluster che attraversano i filtri.
	Align 1 (Allineamento 1)	Percentuale di cluster allineati al riferimento nella Read 1 (Lettura 1).
	Align 2 (Allineamento 2)	Percentuale di cluster allineati al riferimento nella Read 2 (Lettura 2).
	I20 / I1 1	Rapporto di intensità al ciclo 20 rispetto alle intensità al ciclo 1 per Read 1 (Lettura 1).
	I20 / I2 1	Rapporto di intensità al ciclo 20 rispetto alle intensità al ciclo 1 per Read 2 (Lettura 2).
	PE Resynthesis (Risintesi PE)	Rapporto fra le intensità del primo ciclo per Read 1 (Lettura 1) e le intensità del primo ciclo per Read 2 (Lettura 2).

Grafico Cluster (Cluster)

Asse Y	Asse X	Descrizione
Clusters (Cluster)	Raw (Non elaborato)	Numero totale dei cluster rivelati durante la corsa.
	PF (Che attraversano il filtro)	Numero totale dei cluster che attraversano il filtro durante la corsa.
	Unaligned (Non allineato)	Numero totale di cluster che attraversano il filtro non allineati al genoma di riferimento, se applicabile. I cluster non indicizzati non sono inclusi nella conta dei non allineati.
	Unindexed (Non indicizzato)	Numero totale dei cluster che attraversano il filtro non associati ad alcuna sequenza d'indice durante la corsa.
	Duplicate (Duplicato)	Questo valore non è applicabile ai saggi per fibrosi cistica o a Universal Kit 1.0 e sarà sempre uguale a zero.

Grafico Mismatch (Mancata corrispondenza)

Asse Y	Asse X	Descrizione
Percent (Percentuale)	Cycle (Ciclo)	Rappresenta graficamente la percentuale delle mancate corrispondenze per tutti i cluster in una corsa per ciascun ciclo.

Scheda Details (Dettagli) per il saggio CF 139-Variant

Le informazioni visualizzate nella scheda Details (Dettagli) per il saggio CF 139-Variant comprendono una tabella dei campioni e una tabella delle varianti.

Figura 14 Esempio di scheda Details (Dettagli) per il saggio CF 139-Variant

#	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	PolyTGT/PolyT	PolyTGT/PolyT	N/A	Intron 9	117188661-117188689	c.1210-12T15_9]	N/A	(TG)10(T)7/(TG)10(T)9	53
2	NA01445	NA01445	F508del	DEL	rs113993860	Exon 11	117199645	c.1521_1523delCTT	p.Phe508del	HET	307
3	NA01445	NA01445	W1282X	SNV	rs77010898	Exon 23	117282820	c.3846G>A	p.Trp1282X	HET	307

Tabella Samples (Campioni) per il saggio CF 139-Variant

Colonna	Descrizione
# (N.)	Numero ordinale di identificazione all'interno della tabella.
Sample ID (ID campione)	ID del campione nel foglio campioni. L'ID campione deve essere sempre un valore univoco.
Sample Name (Nome del campione)	Nome del campione nel foglio campioni.
Call Rate (Percentuale di identificazione)	Numero di posizioni delle mutazioni che soddisfano un valore di soglia di affidabilità predefinito diviso il numero totale di posizioni delle mutazioni interrogate. La percentuale di identificazione viene descritta in base al singolo campione e riportata come percentuale calcolata come 1 meno [numero di posizioni con identificazioni incomplete diviso il numero totale di posizioni sequenziate].

Colonna	Descrizione
Performance (Prestazioni)	Frequenza di Pass (Superato) o Fail (Non superato) in base alla percentuale di identificazione. Per un campione di controllo positivo: <ul style="list-style-type: none"> • PASS: con una percentuale di identificazione $\geq 99\%$ • FAIL (Non superato): con una percentuale di identificazione $< 99\%$ Per un campione di controllo negativo: <ul style="list-style-type: none"> • PASS: con una percentuale di identificazione $\leq 10\%$ • FAIL (Non superato): con una percentuale di identificazione $> 10\%$ Per un campione non marcato come controllo positivo o negativo: <ul style="list-style-type: none"> • PASS: con una percentuale di identificazione $\geq 99\%$ • FAIL (Non superato): con una percentuale di identificazione $< 99\%$
Control (Controllo)	Tipo di controllo in base all'elenco nel foglio campioni. I valori sono positivi o negativi. Un campo vuoto indica solo campione.
Comment (Commento)	Campo di testo opzionale per i commenti. I commenti inseriti in questo campo vengono salvati nel file report dell'analisi, MiSeqDxCf139VariantAssay.txt. Se l'analisi viene rimessa in coda, viene scritto un nuovo file report. I commenti della corsa di un'analisi precedente non vengono riportati nella corsa dell'analisi successiva.

Tabella Variants (Varianti) per il saggio CF 139-Variant

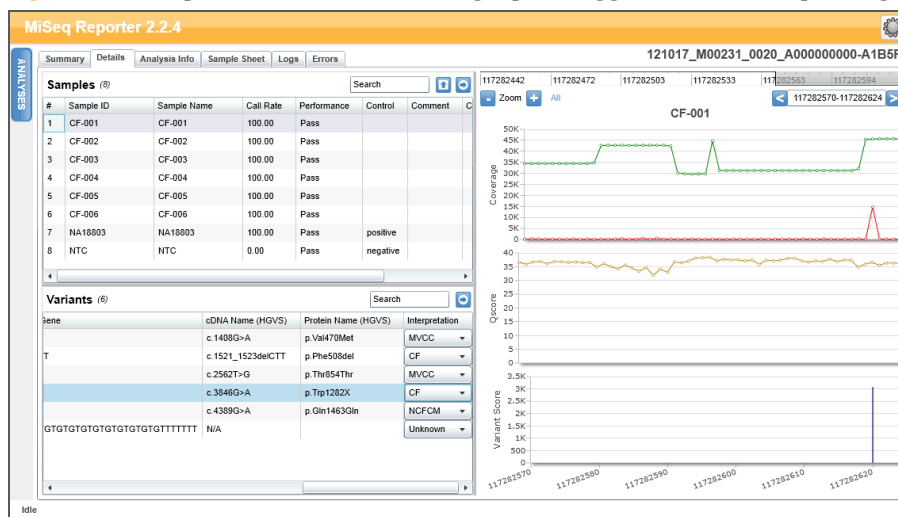
Colonna	Descrizione
# (N.)	Numero ordinale di identificazione all'interno della tabella.
Sample ID (ID campione)	ID del campione nel foglio campioni. L'ID campione deve essere sempre un valore univoco.
Sample Name (Nome del campione)	Nome del campione nel foglio campioni.
Mutations (Common Name) [Mutazioni (Nome comune)]	Nome comune della variante di fibrosi cistica descritta nel database CFTR2.
Mutation Type (Tipo di mutazione)	Tipo di variante. <ul style="list-style-type: none"> • SNV (Single Nucleotide Variant): variante di singolo nucleotide • DIV (Deletion Insertion Variant): variante di delezione-inserzione • DEL (Large deletion): ampia delezione • PolyTGPolyT: genotipo poliTG/poliT nel gene della CF
dbSNP rsID (ID dbSNP rs)	ID dbSNP rs della variante, se applicabile
CFTR Gene Region (Regione del gene CFTR)	Regione del gene CFTR (n. esone o n. introne) in cui è presente la variante.

Colonna	Descrizione
Genomic Location (Posizione genomica)	Posizione genomica della variante.
cDNA Name (HGVS) (Nome cDNA - HGVS)	Descrizione di una variante a livello di DNA mediante la nomenclatura della sequenza di DNA codificante (cDNA) raccomandata dalla Human Genome Variation Society (HGVS).
Protein Name (HGVS) (Nome proteina - HGVS)	Descrizione di una variante a livello di proteina mediante la nomenclatura della sequenza proteica raccomandata dalla Human Genome Variation Society (HGVS).
Result (Risultato)	Genotipo della variante. Per le varianti SNV, DIV e DEL: <ul style="list-style-type: none"> HET (Heterozygous): eterozigote HOM (Homozygous): omozigote Per le varianti poliTG/poliT, viene riportato il genotipo effettivo. NOTA: PolyTG/PolyT è riportato solo quando viene rilevata la variante R117H.

Scheda Details (Dettagli) per il saggio CF Clinical Sequencing

Le informazioni visualizzate nella scheda Details (Dettagli) per il saggio CF Clinical Sequencing comprendono una tabella dei campioni, una tabella delle varianti, un grafico della copertura, un grafico del punteggio qualitativo e un grafico del punteggio delle varianti.

Figura 15 Esempio di scheda Details (Dettagli) per il saggio CF Clinical Sequencing



Scheda Samples (Campioni) per il saggio CF Clinical Sequencing

Colonna	Descrizione
# (N.)	Numero ordinale di identificazione all'interno della tabella.
Sample ID (ID campione)	ID del campione nel foglio campioni. L'ID campione deve essere sempre un valore univoco.
Sample Name (Nome del campione)	Nome del campione nel foglio campioni.
Call Rate (Percentuale di identificazione)	Numero di basi che soddisfa una soglia di punteggio qualitativo diviso per il numero totale di basi interrogate. La percentuale di identificazione viene descritta in base al singolo campione ed espressa in percentuale calcolata come 1 meno [numero di posizioni con identificazioni incomplete diviso il numero totale di basi/posizioni sequenziate].
Performance (Prestazioni)	Frequenza di Pass (Superato) o Fail (Non superato) in base alla percentuale di identificazione. Per un campione di controllo positivo: <ul style="list-style-type: none"> • PASS: con una percentuale di identificazione $\geq 99\%$ • FAIL (Non superato): con una percentuale di identificazione $< 99\%$ Per un campione di controllo negativo: <ul style="list-style-type: none"> • PASS: con una percentuale di identificazione $\leq 10\%$ • FAIL (Non superato): con una percentuale di identificazione $> 10\%$ Per un campione non marcato come controllo positivo o negativo: <ul style="list-style-type: none"> • PASS: con una percentuale di identificazione $\geq 99\%$ • FAIL (Non superato): con una percentuale di identificazione $< 99\%$
Control (Controllo)	Tipo di controllo in base all'elenco nel foglio campioni. I valori sono positivi o negativi. Un campo vuoto indica solo campione.
Comment (Commento)	Campo di testo opzionale per i commenti. I commenti inseriti in questo campo vengono salvati nel file report dell'analisi, MiSeqDxCFClinicalSequencing.txt. Se l'analisi viene rimessa in coda, viene scritto un nuovo file report. I commenti della corsa di un'analisi precedente non vengono riportati nella corsa dell'analisi successiva.
Coordinates Not Called (Coordinate non identificate)	Coordinate genomiche nella regione mirata in cui un'identificazione non è stata riportata a causa di bassi valori di affidabilità.

Scheda Variants (Varianti) per il saggio CF Clinical Sequencing

Colonna	Descrizione
# (N.)	Numero ordinale di identificazione all'interno della tabella.

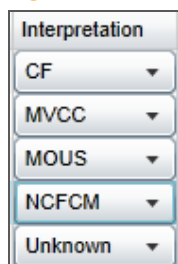
Colonna	Descrizione
Sample ID (ID campione)	ID del campione nel foglio campioni. L'ID campione deve essere sempre un valore univoco.
Sample Name (Nome del campione)	Nome del campione nel foglio campioni.
Chr (Crom)	Obiettivo di riferimento o nome del cromosoma.
Position (Posizione)	Posizione in cui si è stata trovata la variante.
Variant Type (Tipo di variante)	Tipo di variante. <ul style="list-style-type: none"> • SNV (Single Nucleotide Variant): variante di singolo nucleotide • DIV (Deletion Insertion Variant): variante di delezione-inserzione • DEL (Large deletion): ampia delezione • PolyTGPolyT: genotipo poliTG/poliT nel gene della CF
Call (Identificazione)	Stringa che rappresenta come è cambiata la base o sono cambiate le basi in questa posizione nel riferimento.
Frequency (Frequenza)	Frazione di letture del campione che include la variante. Per esempio, se la base di riferimento in una particolare posizione è A e il campione 1 ha 60 letture A e 40 letture T, allora la SNV ha una frequenza di variante pari a 0,4.
Depth (Profondità)	Numero di letture per un campione che copre una particolare posizione.
Filter (Filtro)	Criterio per una variante filtrata.
dbSNP ID	Nome dbSNP della variante.
RefGene	Gene secondo il RefGene in cui appare questa variante.
cDNA Name (HGVS) (Nome cDNA - HGVS)	Descrizione di una variante a livello di DNA mediante la nomenclatura della sequenza di DNA codificante (cDNA) raccomandata dalla Human Genome Variation Society (HGVS).
Protein Name (HGVS) (Nome proteina - HGVS)	Descrizione di una variante a livello di proteina mediante la nomenclatura della sequenza proteica raccomandata dalla Human Genome Variation Society (HGVS).
Interpretation (Interpretazione)	Questo campo consente al genetista di fornire un'interpretazione clinica della mutazione di ciascun campione. L'elenco a discesa comprende le seguenti opzioni per ciascun campione: <ul style="list-style-type: none"> • CF (Cystic fibrosis): che causa la fibrosi cistica • MVCC (Mutation of Varying Clinical Consequence): mutazione di varie conseguenze cliniche • MOUS (Mutation of Unknown Significance): mutazione con significato sconosciuto • NCFM (Non CF Causing Mutation): mutazione che non causa la fibrosi cistica • Unknown (Sconosciuta) Mediante l'icona è possibile generare un nuovo report.

Colonna Interpretation (Interpretazione) della tabella delle varianti

Nella colonna Interpretation (Interpretazione) sono riportate le selezioni che consentono al genetista di interpretare le mutazioni di ciascun campione. L'elenco a discesa Interpretation (Interpretazione) comprende le seguenti opzioni:

- **CF** (Cystic fibrosis): che causa la fibrosi cistica
- **MVCC** (Mutation of Varying Clinical Consequence): mutazione di varie conseguenze cliniche
- **MOUS** (Mutation of Unknown Significance): mutazione con significato sconosciuto
- **NCFCM** (Non CF Causing Mutation): mutazione che non causa la fibrosi cistica
- **Unknown** (Sconosciuta)

Figura 16 Colonna Interpretation (Interpretazione)



I risultati delle tabelle Variants (Varianti) possono essere esportati singolarmente in un file di testo mediante l'icona **Export table data to text file** (Esporta dati della tabella in un file di testo). L'esportazione non altera il file report dell'analisi.



Una volta che il genetista ha completato la determinazione del significato delle varianti, è possibile salvare le impostazioni dell'interpretazione nel report di analisi. Al nome del file report dell'analisi originale sarà automaticamente aggiunto un timbro ora/data.

Grafico Coverage (Copertura) per il saggio CF Clinical Sequencing

Asse Y	Asse X	Descrizione
Coverage (Copertura)	Position (Posizione)	La curva verde rappresenta il numero di letture allineate che coprono ciascuna posizione nel riferimento. La curva rossa rappresenta il numero di letture allineate che presentano una identificazione errata in detta posizione nel riferimento. Le varianti SNV e altre sono rappresentate come aggiunte nella curva rossa.

Grafico Qscore (Punteggio qualitativo)

Asse Y	Asse X	Descrizione
Qscore (Punteggio qualitativo)	Position (Posizione)	Punteggio qualitativo medio delle basi nella posizione del riferimento indicata.

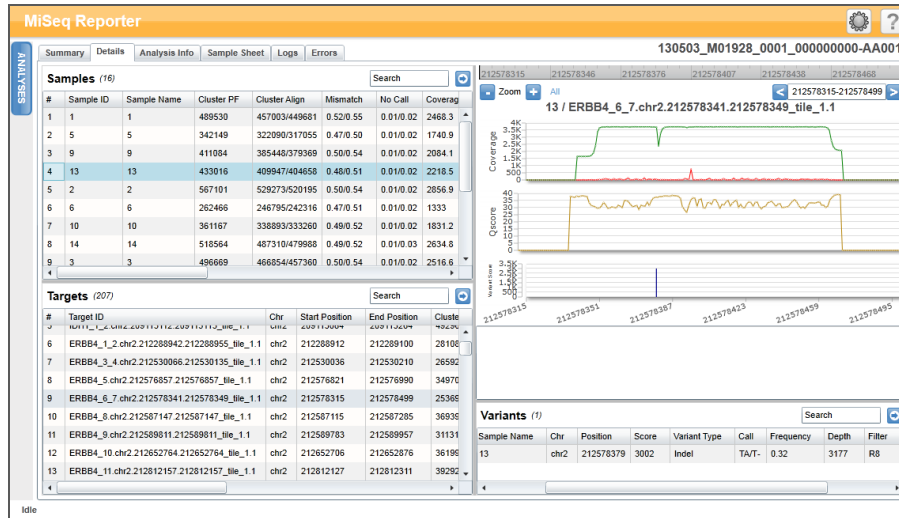
Grafico Variant Score (Punteggio varianti) per il saggio CF Clinical Sequencing

Asse Y	Asse X	Descrizione
Score (Punteggio)	Position (Posizione)	Rappresenta graficamente il punteggio qualitativo e la posizione delle SNV e delle InDel.

Scheda Details (Dettagli) per Universal Kit 1.0

Le informazioni visualizzate nella scheda Details (Dettagli) per Universal Kit 1.0 comprendono una tabella dei campioni, una tabella dei target, un grafico della copertura, un grafico del punteggio qualitativo, un grafico del punteggio delle varianti e una tabella delle varianti.

Figura 17 Esempio di scheda Details (Dettagli) per Universal Kit 1.0



Scheda Samples (Campioni) per Universal Kit 1.0

Colonna	Descrizione
# (N.)	Numero ordinale di identificazione all'interno della tabella.
Sample ID (ID campione)	ID del campione nel foglio campioni. L'ID campione deve essere sempre un valore univoco.
Sample Name (Nome del campione)	Nome del campione nel foglio campioni.
Cluster PF (Cluster che attraversano il filtro)	Il numero di cluster che attraversano il filtro per il campione.
Cluster Align (Allineamento cluster)	Il conteggio totale di cluster che attraversano il filtro che si allineano per il campione (Read 1/Read 2 - Lettura 1/Lettura 2).
Mismatch (Mancata corrispondenza)	La percentuale di mancata corrispondenza sul riferimento e sottoposta a media sui cicli per lettura (Read 1/Read 2 - Lettura 1/Lettura 2).

Colonna	Descrizione
No Call (Nessuna identificazione)	La percentuale di basi che non sono state identificate (mancata identificazione) per il campione e sottoposta a media per i cicli per lettura (Read 1/Read 2 - Lettura 1/Lettura 2).
Coverage (Copertura)	Copertura media (numero di basi allineate su una data posizione del riferimento) sottoposta a media su tutte le posizioni.
Het SNPs (SNP eterozigoti)	Il numero di SNP eterozigoti rilevati per il campione.
Hom SNPs (SNP omozigoti)	Il numero di SNP omozigoti rilevati per il campione.
Insertions (Inserzioni)	Il numero di inserzioni rilevate per il campione.
Deletions (Delezioni)	Il numero di delezioni rilevate per il campione.
Manifest (File manifest)	Il file che specifica un genoma di riferimento e le regioni di riferimento target da usare nella fase di allineamento.
Genome (Genoma)	Il nome del genoma di riferimento.

Tabella Targets (Target) per Universal Kit 1.0

Colonna	Descrizione
# (N.)	Numero ordinale di identificazione all'interno della tabella.
Target ID (ID del target)	Il nome del target nel file manifest.
Chr (Crom)	Obiettivo di riferimento o nome del cromosoma.
Start Position (Posizione di avvio)	La posizione di avvio della regione target.
End Position (Posizione finale)	La posizione finale della regione target.
Cluster PF (Cluster che attraversano il filtro)	Il numero di cluster che attraversano il filtro per il target visualizzato per lettura (Read 1/Read 2 - Lettura 1/Lettura 2).

Colonna	Descrizione
Mismatch (Mancata corrispondenza)	La percentuale di basi con mancata corrispondenza sul target e sottoposta a media su tutti i cicli, visualizzati per lettura. $\text{Mismatch} = [\text{media}(\text{errori conteggio nei cicli}) / \text{cluster PF}] * 100$.
No Call (Nessuna identificazione)	La percentuale di basi con nessuna identificazione per il target e sottoposta a media sui cicli, visualizzati per lettura.
Het SNPs (SNP eterozigoti)	Il numero di SNP eterozigoti rilevati per il target su tutti i campioni.
Hom SNPs (SNP omozigoti)	Il numero di SNP omozigoti rilevati per il target su tutti i campioni.
Insertions (Inserzioni)	Il numero di inserzioni rilevate per il target su tutti i campioni.
Deletions (Delezioni)	Il numero di delezioni rilevate per il target su tutti i campioni.
Manifest (File manifest)	Il file che specifica un genoma di riferimento e le regioni di riferimento target da usare nella fase di allineamento.

Grafico Coverage (Copertura) per Universal Kit 1.0

Asse Y	Asse X	Descrizione
Coverage (Copertura)	Position (Posizione)	La curva verde rappresenta il numero di letture allineate che coprono ciascuna posizione nel riferimento. La curva rossa rappresenta il numero di letture allineate che presentano una identificazione errata in detta posizione nel riferimento. I polimorfismi SNP e altre varianti sono rappresentati come aggiunte nella curva rossa.

Grafico Qscore (Punteggio qualitativo)

Asse Y	Asse X	Descrizione
Qscore (Punteggio qualitativo)	Position (Posizione)	Punteggio qualitativo medio delle basi nella posizione del riferimento indicata.

Grafico Variant Score (Punteggio varianti) per Universal Kit 1.0

Asse Y	Asse X	Descrizione
Score (Punteggio)	Position (Posizione)	Rappresenta graficamente il punteggio qualitativo delle varianti e la posizione di SNP e Indel.

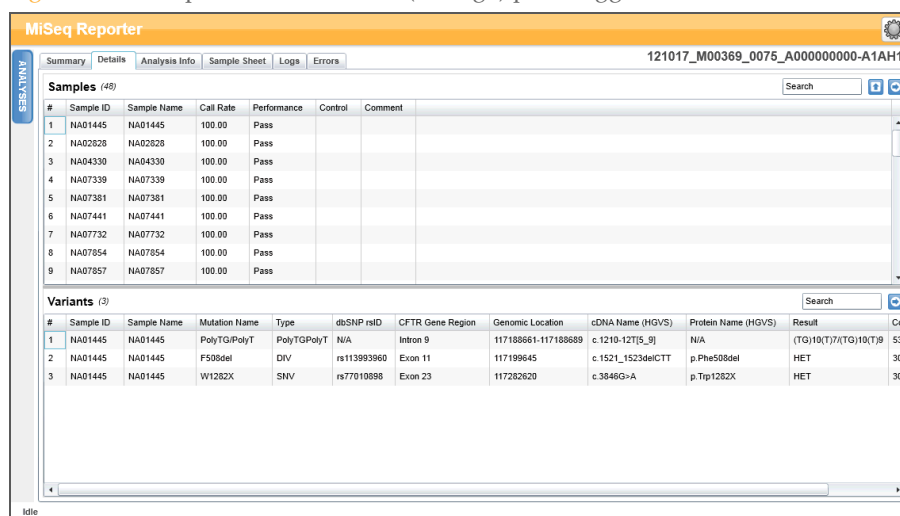
Scheda Variants (Varianti) per Universal Kit 1.0

Colonna	Descrizione
# (N.)	Numero ordinale di identificazione all'interno della tabella.
Sample ID (ID campione)	ID del campione nel foglio campioni. L'ID campione deve essere sempre un valore univoco.
Sample Name (Nome del campione)	Nome del campione nel foglio campioni.
Chr (Crom)	Obiettivo di riferimento o nome del cromosoma.
Position (Posizione)	Posizione in cui si è stata trovata la variante.
Score (Punteggio)	Il punteggio qualitativo delle varianti per questa variante.
Variant Type (Tipo di variante)	Il tipo di variante, che può essere SNP o Indel.
Call (Identificazione)	Una rappresentazione di come una base o delle basi sono cambiate in quella posizione nel riferimento. <ul style="list-style-type: none"> • Gli SNP sono elencati nel formato Riferimento > AlleleA/AlleleB. • Le inserzioni sono elencate nel formato Riferimento/Inserzione. G-/GA mostra l'inserzione di A. • Le delezioni sono elencate nel formato Riferimento/Delezione. ACG/A-- mostra la delezione di GG.
Frequency (Frequenza)	Frazione di letture del campione che include la variante. Ad esempio, se la base di riferimento è A e il campione 1 ha 60 letture A e 40 letture T, allora la SNP ha una frequenza di variante pari a 0,4.
Depth (Profondità)	Numero di letture per un campione che copre una particolare posizione.
Filter (Filtro)	Criterio per una variante filtrata. Se vengono attraversati tutti i filtri, PASS (Superato) viene scritto nella colonna Filter (Filtro). Per maggiori informazioni, vedere <i>Intestazioni e annotazioni del file VCF</i> a pagina 49.
dbSNP	Il nome dbSNP della variante, se applicabile
RefGene	Gene secondo il RefGene in cui appare questa variante.

File di output dell'analisi per i saggi CF

I risultati dell'analisi per i saggi CF vengono visualizzati nella scheda Details (Dettagli).

Figura 18 Esempio di scheda Details (Dettagli) per il saggio CF 139-Variant



#	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	Cor
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



I risultati delle tabelle Variants (Varianti) possono essere esportati singolarmente in un file di testo mediante l'icona **Export table data to text file** (Esporta dati della tabella in un file di testo). L'esportazione non altera il file report dell'analisi.



Una volta che il genetista ha completato la determinazione del significato delle varianti, è possibile salvare le impostazioni dell'interpretazione nel report di analisi. Al nome del file report dell'analisi originale sarà automaticamente aggiunto un timbro ora/data.

I file di output per i saggi CF vengono inoltre riepilogati in un file di testo delimitato da tabulazioni a cui è assegnato il nome del saggio utilizzato per la corsa. Questi risultati sono identici a quelli della scheda Details (Dettagli).

- ▶ Per il saggio CF 139-Variant, il nome del file è MiSeqDxCf139VariantAssay.txt.
- ▶ Per il saggio CF Clinical Sequencing, il nome del file è MiSeqDxCfClinicalSequencingAssay.txt.

Una volta completata l'analisi, il file di output viene scritto nella cartella Alignment (Allineamento) relativa alla corsa. Ad esempio:

MiSeqAnalysis\<<RunFolderName>\Data\Intensities\BaseCalls\Alignment

Se l'analisi è stata ripetuta o rimessa in coda, per quella corsa di analisi, viene scritto un nuovo file report in Alignment (Allineamento). Per maggiori informazioni, vedere *Rimessa in coda di un'analisi* a pagina 13.

Il file di output presenta un'intestazione in cui sono riportate le seguenti informazioni sulla corsa:

Intestazione	Descrizione
Test	Descrive il test eseguito.
Run ID (ID corsa)	ID della corsa generato da MOS all'inizio della corsa di sequenziamento.

Intestazione	Descrizione
Run Date (Data della corsa)	Data (GGMMAA) in cui la corsa di sequenziamento ha avuto inizio in MOS.
Analysis Version (Versione analisi)	Questa è la versione MiSeq Reporter che è stata utilizzata per l'analisi.

Figura 19 Esempio di intestazione del file di output del saggio CF 139-Variant

```
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
```

Sotto l'intestazione si trova una sezione di riepilogo per ciascun ID campione in cui ciascun valore è riportato in colonne. Per una descrizione delle colonne, vedere *Scheda Details (Dettagli)* per il saggio CF 139-Variant a pagina 25 e *Scheda Details (Dettagli)* per il saggio CF Clinical Sequencing a pagina 27.



NOTA

L'architettura pipeline che genera i file di output non è identica per i saggi CF e per Universal Kit 1.0. I file di output generati per Universal Kit 1.0 sono file *.bam, file *.vcf e file AmpliconCoverage_M#.tsv. Per maggiori informazioni sui file di output per Universal Kit 1.0, vedere Appendice A File di output dell'analisi per Universal Kit 1.0.

Installazione e risoluzione dei problemi

Requisiti per l'installazione di MiSeq Reporter su un altro computer	38
Installazione di MiSeq Reporter su un altro computer	39
Utilizzo di MiSeq Reporter su un altro computer	41
Risoluzione dei problemi di MiSeq Reporter	42



Requisiti per l'installazione di MiSeq Reporter su un altro computer

L'installazione di una copia di MiSeq Reporter su un altro computer Windows consente di eseguire un'analisi secondaria dei dati di sequenziamento mentre il MiSeqDx esegue una corsa di sequenziamento successiva.

Per ulteriori informazioni, vedere *Installazione di MiSeq Reporter su un altro computer* a pagina 39.

Requisiti del computer

Per l'utilizzo del software MiSeq Reporter sono necessari i seguenti componenti:

- ▶ Sistema operativo Windows a 64 bit (Vista, Windows 7, Windows Server 2008 a 64 bit)
- ▶ RAM da almeno ≥ 8 GB; si consiglia una RAM da ≥ 16 GB
- ▶ ≥ 1 TB di spazio su disco
- ▶ Processore quad core (2,8 Ghz o superiore)
- ▶ Microsoft .NET 4

Browser supportati

MiSeq Reporter può essere visualizzato con i seguenti browser Web:

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

Download e licenze

- 1 È possibile scaricare una seconda copia del software MiSeq Reporter dal sito web di Illumina. È necessario un login MyIllumina.
- 2 Quando richiesto, durante l'installazione, accettare il contratto di licenza con l'utente finale (End-User Licensing Agreement, EULA). Non è necessario un codice di licenza poiché questa copia supplementare è disponibile gratuitamente.

Installazione di MiSeq Reporter su un altro computer

Per installare MiSeq Reporter su un altro computer Windows, impostare innanzitutto l'autorizzazione **Accedi come servizio**, quindi avviare l'installazione guidata. Successivamente, configurare il software in modo che trovi il Repository e il GenomePath appropriati.

Impostazione dell'account utente o di gruppo su Windows 7

I diritti di amministratore sono necessari per configurare gli account utente o di gruppo e attivare l'autorizzazione **Accedi come servizio**. Se necessario, richiedere l'assistenza dell'amministratore della propria struttura.

- 1 Dal menu **Start** di Windows, selezionare **Pannello di controllo**, quindi fare clic su **Sistema e sicurezza**.
- 2 Fare clic su **Strumenti di amministrazione**, quindi fare doppio clic su **Criteri di protezione locali**.
- 3 Nella struttura Impostazioni protezione a sinistra, fare doppio clic su **Criteri locali**, quindi fare clic su **Assegnazione diritti utente**.
- 4 Nel riquadro dei dettagli a destra, fare doppio clic su **Accedi come servizio**.
- 5 Nella finestra di dialogo Proprietà, fare clic su **Aggiungi utente o gruppo**.
- 6 Digitare il nome account dell'utente o del gruppo per il computer. Fare clic su **Controlla nomi** per convalidare l'account.
- 7 Fare clic su **OK** in tutte le finestre di dialogo aperte, quindi chiudere il pannello di controllo.

Per ulteriori informazioni, vedere [technet.microsoft.com/en-us/library/cc739424\(W.S.10\).aspx](http://technet.microsoft.com/en-us/library/cc739424(W.S.10).aspx) sul sito Web di Microsoft.

Installazione guidata di MiSeq Reporter

- 1 Scaricare e decomprimere il pacchetto di installazione di MiSeq Reporter dal sito Web Illumina.
- 2 Fare doppio clic sul file setup.exe.
- 3 Fare clic su **Next** (Avanti) sui messaggi visualizzati dalla creazione guidata per l'installazione.
- 4 Quando richiesto, specificare il nome utente e la password per l'account con autorizzazione **Accedi come servizio**, in base alle impostazioni del passaggio precedente.
- 5 Continuare con i messaggi successivi.

Configurazione di MiSeq Reporter

Per configurare MiSeq Reporter in modo che trovi la cartella della corsa e la cartella del genoma di riferimento, modificare il file di configurazione in un editor di testo, come Blocco note.

- 1 Selezionare la cartella di installazione (per impostazione predefinita C:\Illumina\MiSeq Reporter) e aprire il file MiSeq Reporter.exe.config con un editor di testo.
- 2 Individuare la tag **Repository** e modificare il **valore** nella posizione predefinita dei dati sul computer non integrato sullo strumento.
Esempio:
`<add key="Repository" value="E:\Data\Repository" />`
In alternativa, la posizione può trovarsi in una rete accessibile dal computer non integrato.
- 3 Trovare la tag **GenomePath** (Percorso genoma) e modificare il **valore** con la posizione della cartella contenente i file in formato FASTA con i genomi di riferimento.
Esempio:
`<add key="GenomePath" value="E:\MyGenomes\FASTA" />`


Avvio del servizio MiSeq Reporter

Una volta completata l'installazione, il servizio MiSeq Reporter si avvia automaticamente. Se questo non accade, avviarlo manualmente attenendosi alle seguenti istruzioni, oppure riavviare il computer.

- 1 Dal menu **Start** di Windows, fare clic con il tasto destro del mouse su **Computer** e selezionare **Gestione**.
- 2 Dalla struttura Gestione computer a sinistra, fare doppio clic su **Servizi e applicazioni**, quindi fare clic su **Servizi**.
- 3 Fare clic con il tasto destro del mouse su **MiSeq Reporter** e selezionare **Proprietà**.
- 4 Nella scheda Generale accertarsi che **Tipo di avvio** sia impostato su **Automatico**, quindi fare clic su **Avvia**.
- 5 Nella scheda Accesso, impostare **nome utente** e **password** per un account Servizi con permessi di scrittura sul server. Illumina consiglia l'account **Sistema locale** per la maggior parte degli utenti. Per assistenza o requisiti di rete specifici del sito, rivolgersi all'amministratore della propria struttura.
- 6 Fare clic su **OK** in tutte le finestre di dialogo aperte, quindi chiudere la finestra Gestione computer.
- 7 Dopo aver avviato il servizio MiSeq Reporter, collegarsi localmente al software inserendo localhost:8042 in un browser Web.

Utilizzo di MiSeq Reporter su un altro computer

Per utilizzare MiSeq Reporter su un altro computer, è necessario che le cartelle contenenti i dati delle corse e i genomi di riferimento siano accessibili.

- 1 Salvo il caso in cui si utilizzi una posizione in rete per i dati di sequenziamento e i genomi di riferimento, copiare le cartelle elencate di seguito nel computer locale:
 - Copiare i dati della corsa dal computer MiSeqDx in
D:\MiSeqOutput*<CartellaCorsa>*.
 - Copiare i genomi di riferimento dal computer MiSeqDx in C:\Illumina\MiSeq Reporter\Genomes.
- 2 Aprire un browser Web all'indirizzo <http://localhost:8042>; si aprirà l'interfaccia Web di MiSeq Reporter.
- 3 Modificare il percorso per il Repository mediante l'icona **Settings** (Impostazioni)  nell'angolo in alto a destra dell'interfaccia Web.



NOTA

Specificando che il percorso del repository in Settings (Impostazioni) è temporaneo, al riavvio del computer, il percorso predefinito per la posizione del repository sarà quello specificato in MiSeq Reporter.exe.config.

- 4 Selezionare **Analyses** (Analisi) nella parte sinistra dell'interfaccia Web per visualizzare le corse disponibili nella posizione specificata del Repository.
- 5 Prima di poter rimettere in coda un'analisi mediante MiSeq Reporter installato su un altro computer, è necessario aggiornare il percorso per la cartella GenomeFolder nel foglio campioni dalla scheda Sample Sheet (Foglio campioni). Una volta aggiornato il percorso GenomeFolder, fare clic su **Save and Requeue** (Salva e rimetti in coda). Per maggiori informazioni, vedere *Modifica del foglio campioni in MiSeq Reporter* a pagina 11.

Risoluzione dei problemi di MiSeq Reporter

MiSeq Reporter opera come applicazione di servizio di Windows. Gli account utente devono essere configurati per attivare le autorizzazioni **Accedi come servizio** prima di installare MiSeq Reporter. Per maggiori informazioni, vedere *Impostazione dell'account utente o di gruppo su Windows 7* a pagina 39.

Per maggiori informazioni, vedere msdn.microsoft.com/en-us/library/ms189964.aspx.


Errore di avvio del servizio

Se il servizio non si avvia, controllare il registro eventi di Windows e visualizzare i dettagli del messaggio di errore.

- 1 Aprire **Pannello di controllo** e selezionare **Strumenti di amministrazione**.
- 2 Selezionare **Visualizzatore eventi**.
- 3 Nella finestra Visualizzatore eventi, selezionare **Registri di Windows | Applicazione**. L'errore elencato nel registro eventi descrive qualsiasi errore di sintassi in MiSeq Reporter.exe.config. Una sintassi errata nel file MiSeq Reporter.exe.config può impedire il funzionamento del servizio.

Errore di copia file

Se i file non vengono copiati nella posizione desiderata, verificare le seguenti impostazioni:

- 1 Controllare il percorso della cartella repository specificata o della cartella MiSeqOutput:
 - Per le installazioni su un altro computer, verificare la posizione del repository mediante l'icona Settings (Impostazioni)  sull'interfaccia web di MiSeq Reporter.
 - Per le installazioni integrate sullo strumento, verificare la posizione della cartella MiSeqOutput sulla schermata Run Options (Opzioni corsa), scheda Folder Settings (Impostazioni cartella) di MOS.

Si deve utilizzare il percorso completo UNC (per es. \\server1\Runs. Dato che MiSeq Reporter opera come servizio di Windows, non riconosce le unità mappate dall'utente (per es. Z:\Runs).

- 2 Confermare l'accesso in scrittura alla posizione della cartella di output. Per assistenza, rivolgersi all'amministratore della propria struttura.
- 3 Verificare che l'opzione di copia non sia disattivata in MiSeq Reporter.exe.config. Questa impostazione si trova nella sezione <appSettings> e il valore deve essere impostato su **1**.


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

Visualizzazione di file Log (Registro) per una corsa non riuscita

La visualizzazione di file Log (Registro) può contribuire all'identificazione di errori specifici per la risoluzione dei problemi.

- 1 Per visualizzare i file Log (Registro) con l'interfaccia Web di MiSeq Reporter, selezionare la corsa nella scheda Analyses (Analisi).
- 2 Selezionare la scheda Log (Registro) per visualizzare un elenco di ogni fase di svolgimento dell'analisi. Le informazioni dei file di registro sono registrate nel file AnalysisLog.txt che si trova a livello della radice della cartella MiSeqAnalysis.
- 3 Selezionare la scheda Errors (Errori) per visualizzare un elenco di errori verificatisi durante l'analisi. Le informazioni sugli errori sono registrate nel file AnalysisError.txt che si trova a livello della radice della cartella MiSeqAnalysis.

[Questa pagina è stata lasciata intenzionalmente vuota]

File di output dell'analisi per Universal Kit 1.0

Tipi di file di output dell'analisi	46
Formato file BAM	47
Formato file VCF	48
File di copertura dell'amplicone	51
File di output supplementari	52



Tipi di file di output dell'analisi

La seguente tabella descrive i file di output generati per Universal Kit 1.0, che forniscono i risultati dell'analisi per l'allineamento, l'identificazione delle varianti e la copertura.

Nome file	Descrizione
File *.bam	Contiene le letture allineate per un dato campione. Si trova in Data\Intensities\BaseCalls\Alignment.
File *.vcf	Contiene le informazioni sulle varianti identificate nelle posizioni specifiche in un genoma di riferimento. Si trova in Data\Intensities\BaseCalls\Alignment.
AmpliconCoverage_M#.tsv	Contiene i dettagli sulla copertura risultante per amplicone per campione. M# rappresenta il numero di manifest. Si trova in Data\Intensities\BaseCalls\Alignment.



NOTA

L'architettura pipeline che genera questi file di output non è identica per i saggi CF e per Universal Kit 1.0. Questa sezione descrive i file di output dell'analisi solo per Universal Kit 1.0.

Formato file BAM

Un file BAM (*.bam) è una versione binaria compressa di un file SAM usato nella rappresentazione delle sequenze allineate. I formati SAM e BAM sono descritti nei dettagli sul sito Web SAM Tools: samtools.sourceforge.net.

I file BAM sono scritti nella cartella di allineamento in Data\Intensities\BaseCalls\Alignment nel formato di nome del file NomeCampione_S#.bam dove # rappresenta il numero di campioni determinati dall'ordine in cui i campioni sono elencati nel foglio campioni.

I file BAM contengono una sezione di intestazione e una sezione di allineamenti:

- ▶ **Intestazione:** contiene le informazioni sull'intero file, come il nome del campione e la lunghezza del campione. Gli allineamenti nella sezione degli allineamenti sono associati con le informazioni specifiche contenute nella sezione di intestazione.
- ▶ **Allineamenti:** contiene il nome della lettura, la sequenza della lettura, la qualità della lettura e le tag personalizzate.

Figura 20 Esempio di sezione di allineamento del file 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
```

Il nome della lettura include il cromosoma e le coordinate di avvio (**chr22 17552189**), la qualità dell'allineamento (**8**) e il descrittore della corrispondenza (**35M * 0 0**).

I file BAM sono adatti per la visualizzazione con un visualizzatore esterno come IGV o UCSC Genome Browser.

Formato file VCF

VCF è un formato file ampiamente usato, sviluppato dalla comunità scientifica genomica che contiene informazioni sulle varianti identificate in posizioni specifiche in un genoma di riferimento.

I file VCF usano un formato di denominazione dei file NomeCampione_S#.vcf dove # è il numero del campione in base all'ordine in cui i campioni sono elencati nel foglio campioni.

- ▶ **Intestazione del file VCF:** include la versione del formato file VCF e la versione di Variant Caller. L'intestazione elenca le annotazioni usate nel resto del file. L'ultima riga nell'intestazione contiene le intestazioni delle colonne per le righe dei dati. Per maggiori informazioni, vedere *Intestazioni e annotazioni del file VCF* a pagina 49.

Figura 21 Esempio di intestazione di file 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
```

- ▶ **Righe dei dati del file VCF:** contiene informazioni su una singola variante. Le righe dei dati sono elencate sotto le intestazioni delle colonne incluse nell'intestazione.

Intestazioni e annotazioni del file VCF

Il formato file VCF è flessibile e ampliabile. Le tabelle seguenti descrivono le intestazioni e le annotazioni del file VCF generate da MiSeq Reporter.

Intestazioni del file VCF

Intestazione	Descrizione
CHROM (Cromosoma)	Il cromosoma del genoma di riferimento. I cromosomi appaiono nello stesso ordine del file di riferimento FASTA.
POS (Posizione)	La posizione di singola base della variante nel cromosoma di riferimento. Per gli SNP, questa posizione è la base di riferimento con la variante; per gli Indel o le delezioni, questa posizione è la base di riferimento immediatamente prima della variante.
ID (Identificazione)	Il numero rs per gli SNP ottenuti da dbSNP.txt, se applicabile. Se sono presenti numeri rs multipli in questa posizione, l'elenco è delimitato da punti e virgole. Se non esistono voci dbSNP in questa posizione, viene usato un indicatore di valore mancante ('.').
REF (Riferimento)	Il genotipo di riferimento. Ad esempio, una delezione di T singolo è rappresentata come TT di riferimento e T alternato.
ALT (Alternato)	Gli alleli diversi dalla lettura di riferimento. Ad esempio, un'inserzione di T singolo è rappresentata come A di riferimento e AT alternata.
QUAL (Qualità)	Un punteggio qualitativo su scala Phred assegnato da Variant Caller. Punteggi elevati indicano un'affidabilità superiore nella variante e minore probabilità di errori. Per un punteggio qualitativo di Q, la probabilità di errore stimata è $10^{-(Q/10)}$. Ad esempio, il set di identificazioni con punteggio qualitativo Q30 ha una percentuale di errore di 0,1%. Diversi Variant Caller assegnano punteggi qualitativi in base ai propri modelli statistici, che sono molto relativi alla percentuale di errori osservata.

Annotazioni del file VCF

Intestazione	Descrizione
FILTER (Filtro)	Se vengono attraversati tutti i filtri, PASS (Superato) viene scritto nella colonna Filter (Filtro). <ul style="list-style-type: none"> • LowDP (Profondità bassa): applicato ai siti con profondità di copertura sotto il valore di cutoff. • LowGQ (Qualità genotipizzazione bassa): la qualità di genotipizzazione (GQ) è sotto il valore di cutoff. • LowQual (Qualità bassa): la qualità della variante (QUAL) è sotto il valore di cutoff. • LowVariantFreq (Frequenza varianti bassa): la frequenza delle varianti è inferiore alla soglia. • R8: per gli indel, il numero di ripetizioni adiacenti (base 1 o base 2) nel riferimento è superiore a 8.

Intestazione	Descrizione
INFO (Informazioni)	<ul style="list-style-type: none"> • AC (Conteggio alleli): conteggio degli alleli nei genotipi per ciascun allele ALT, nello stesso ordine in cui sono elencati. • AF (Frequenza allelica): frequenza allelica per ciascun allele ALT, nello stesso ordine in cui sono elencati. • AN (Numero alleli) : il numero totale di alleli nei genotipi identificati. • Exon (Esone): un elenco separato da virgola delle regioni esoniche lette da RefGene. • FC (Conseguenze funzionali): conseguenze funzionali. • GI (Geni identificati): un elenco separato da virgola degli ID del gene letti da RefGene. • TI (Trascritti identificati): un elenco separato da virgola degli ID del trascritto letti da RefGene.
FORMAT (Formato)	<ul style="list-style-type: none"> • AD (Assi): voci nel formato X,Y, dove X rappresenta il numero delle identificazioni di riferimento e Y il numero di identificazioni alternate. • DP (Profondità): profondità approssimativa della lettura; letture con MQ=255 o con accoppiamenti non corretti sono filtrate. • GQ (Qualità genotipo): qualità del genotipo. • GQX (Qualità genotipo X): qualità del genotipo. GQX rappresenta il minimo del valore GQ (Qualità genotipo) e la colonna QUAL (Qualità). In generale, questi valori sono simili; se si prende il valore minimo, GQX diventa la misura più conservativa della qualità del genotipo. • GT (Genotipo): il genotipo. 0 corrisponde alla base di riferimento, 1 corrisponde alla prima voce nella colonna ALT (Alternato), e così via. Il simbolo slash in avanti (/) indica che non è disponibile alcuna informazioni sulla fase. • VF (Frequenza variante): frequenza della variante; la percentuale di letture che supportano l'allele alternato.
SAMPLE (Campione)	La colonna dei campioni fornisce il valore specificato nella colonna FORMAT (Formato).

File di copertura dell'amplicone

Per ciascun file manifest viene generato un file di copertura dell'amplicone. M# nel nome del file rappresenta il numero di file manifest come elencato nel foglio campioni.

Ciascun file inizia con una riga di intestazione che contiene l'ID campione associato con il file manifest. La prima colonna contiene l'ID del target. Ciascuna colonna aggiuntiva elenca la profondità di copertura per l'ID campione associato.

File di output supplementari

I file di output seguenti forniscono informazioni supplementari o riassumono i risultati della corsa e gli errori dell'analisi. Sebbene questi file non siano richiesti per valutare i risultati dell'analisi, possono essere usati per la risoluzione dei problemi.

Nome file	Descrizione
AnalysisLog.txt	Il registro dell'elaborazione che descrive tutte le fasi che si sono verificate durante l'analisi della cartella della corsa attuale. Questo file non contiene messaggi di errore. Si trova a livello della radice della cartella della corsa.
AnalysisError.txt	Il registro dell'elaborazione che elenca qualsiasi errore verificatosi durante l'analisi. Questo file è presente solo se si verifica un errore. Si trova a livello della radice della cartella della corsa.
AmpliconRunStatistics.xml	Contiene un riepilogo delle statistiche specifiche per la corsa. Si trova a livello della radice della cartella della corsa.
CompletedJobInfo.xml	Scritto dopo il completamento dell'analisi, contiene informazioni sulla corsa, come data, ID della cella a flusso, versione software e altri parametri. Si trova a livello della radice della cartella della corsa.
DemultiplexSummaryF1L1.txt	Report sui risultati di demultiplex in una tabella con una riga per tile e una colonna per campione. Si trova in Data\Intensities\BaseCalls\Alignment.
ErrorsAndNoCallsByLaneTileReadCycle.csv	Un file con valori separati da virgola che contiene la percentuale di errore e le identificazioni non riuscite per ciascuna tile, lettura e ciclo. Si trova in Data\Intensities\BaseCalls\Alignment.
Mismatch.htm	Contiene gli istogrammi delle mancate corrispondenze per ciclo e identificazioni non riuscite per ciclo per ciascuna tile. Si trova in Data\Intensities\BaseCalls\Alignment.
Summary.xml	Contiene un riepilogo delle percentuali di mancata corrispondenza e altri risultati dell'identificazione delle basi. Si trova in Data\Intensities\BaseCalls\Alignment.
Summary.htm	Contiene un riepilogo della pagina Web generato da Summary.xml. Si trova in Data\Intensities\BaseCalls\Alignment.

*

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

A

accesso a un servizio 39
 Account Sistema locale 40
 allineamento 16
 analisi
 durante il sequenziamento 2
 AnalysisError.txt 43
 AnalysisLog.txt 43
 annotazioni del file VCF 49
 assistenza clienti 55
 assistenza tecnica 55
 avvio del servizio non riuscito 42

C

cartella analisi 9, 17
 cartella copia 9
 cartella corsa
 rapporto 17
 cartella corse
 informazioni su 4
 cartella dati 9
 cartella MiSeqAnalysis 17
 cartella MiSeqOutput 17
 che attraversano il filtro (PF) 14
 cicli lettura 9
 cluster che attraversano il filtro 14

D

database dbsnp 21
 database refGene 21
 database, preinstallate 21
 demultiplex 16
 determinazione fasi (phasing),
 predeterminazione fasi
 (prephasing) 15
 DLSO 22
 documentazione 55

F

file BAM
 formato file 47
 file di input 21
 file FASTQ 16
 file indice BAM 47
 file Log (Registro) 43
 file manifest 4, 10
 file VCF
 formato file 48
 flusso di lavoro
 Custom Amplicon (Amplicone
 personalizzato) 22

software MiSeqDx 2
 flusso di lavoro Custom Amplicon
 (Amplicone personalizzato) 22
 foglio campioni
 informazioni su 4
 modifica 11
 frequenza varianti VF 49

G

genomi di riferimento, preinstallati 21
 genotipo GT 49
 grafico Clusters (Cluster) 7
 grafico Coverage (Copertura) 8
 grafico High Percentages (Percentuali
 alte) 7
 grafico Low Percentages (Percentuali
 basse) 7
 grafico Mismatch (Mancata
 corrispondenza) 7
 grafico Qscore (Punteggio qualitativo) 8
 grafico Variant Score (Punteggio
 varianti) 8
 guida, tecnica 55

I

icone, stato dell'analisi 6
 ID gene GI 49
 ID trascritto TI 49
 impossibile copiare i file 42
 indirizzo IP, MiSeq Reporter 3
 installazione 39

K

kit 20

L

licenza (EULA) 38
 localhost 3
 LowDP (Profondità copertura bassa) 49
 LowGQ (Qualità genotipizzazione
 bassa) 49
 LowVariantFreq (Frequenza varianti
 bassa) 49

M

MiSeqDxCf139VariantAssay.txt 35
 MiSeqDxCfClinicalSequencingAssay.txt
 35
 modifica foglio campioni 11

P

percorso genoma 39
 percorso repository 5, 39
 policy di sicurezza locale 39
 probabilità di errore 14

punteggi qualitativi 14
punteggi qualitativi (Q-scores) 14

R

r8 49
requisiti del computer 38
rimessa in coda analisi 13
rimessa in coda dell'analisi 6, 11
risoluzione dei problemi
 avvio del servizio non riuscito 42
 file Log (Registro) 43
 impossibile copiare i file 42
RTAComplete.txt 21
RunInfo.xml 21

S

saggi 20
saggio CF 139-Variant 20
saggio CF Clinical Sequencing 20
SAM Tools 47
SampleSheet.csv 21
servizio Windows
 Accedi come servizio 42
 informazioni 2
Smith-Waterman 16, 22
su un altro computer 39

T

tabella Samples (Campioni) 8
tabella Variants (Varianti) 8

U

ULSO 22
Universal Kit 1.0 20
URL server 5

V

visualizzazione di MiSeq Reporter 3

Assistenza tecnica

Per l'assistenza tecnica, contattare l'Assistenza tecnica Illumina.

Tabella 4 Dati di contatto generali Illumina

Sito Web	www.illumina.com
E-mail	techsupport@illumina.com

Tabella 5 Numeri di telefono dell'Assistenza clienti Illumina

Area geografica	Numero di contatto	Area geografica	Numero di contatto
Nord America	1.800.809.4566	Italia	800.874909
Australia	1.800.775.688	Norvegia	800.16836
Austria	0800.296575	Nuova Zelanda	0800.451.650
Belgio	0800.81102	Paesi Bassi	0800.0223859
Cina	400.635.9898	Regno Unito	0800.917.0041
Danimarca	80882346	Singapore	1.800.579.2745
Finlandia	0800.918363	Spagna	900.812168
Francia	0800.911850	Svezia	020790181
Germania	0800.180.8994	Svizzera	0800.563118
Giappone	0800.111.5011	Taiwan	00806651752
Hong Kong	800960230	Altri paesi	+44.1799.534000
Irlanda	1.800.812949		

Schede di sicurezza (SDS)

Le schede di sicurezza (Safety Data Sheet, SDS) sono disponibili sul sito Web Illumina all'indirizzo support.illumina.com/sds.html.

Documentazione dei prodotti

La documentazione dei prodotti in formato PDF può essere scaricata dal sito web Illumina. Andare alla pagina support.illumina.com, selezionare un prodotto, quindi fare clic su **Documentation & Literature** (Documentazione e letteratura).

