

## Inserto della confezione degli strumenti con MOS v4

PER USO DIAGNOSTICO IN VITRO. SOLO PER L'ESPORTAZIONE.

## Uso previsto

Lo strumento MiSeqDx è destinato al sequenziamento mirato di librerie di DNA ottenute da DNA genomico umano estratto da sangue intero periferico o da tessuto fissato in formalina e incluso in paraffina (Formalin-Fixed, Paraffin-Embedded, FFPE), nell'ambito di saggi diagnostici *in vitro* (In Vitro Diagnostic, IVD) eseguiti con lo strumento. Lo strumento MiSeqDx non è destinato al sequenziamento dell'intero genoma o al sequenziamento *de novo*. Lo strumento MiSeqDx deve essere utilizzato con reagenti IVD registrati ed elencati, autorizzati o approvati e software analitico.

## Principi della procedura

Lo strumento Illumina MiSeqDx è destinato al risequenziamento mirato di DNA umano utilizzando i materiali di consumo per il sequenziamento Illumina e le librerie preparate da DNA genomico umano estratto da sangue intero periferico o tessuto FFPE utilizzando reagenti IVD registrati ed elencati, autorizzati o approvati. Le librerie sono preparate amplificando i target e aggiungendo gli indici campione e le sequenze di cattura. Le librerie di campioni sono catturate su una cella a flusso e sequenziate sullo strumento utilizzando la chimica di sequenziamento mediante sintesi (SBS). La chimica SBS utilizza un metodo che fa uso di terminatori reversibili per rilevare le singole basi nucleotidiche man mano che vengono incorporate in filamenti di DNA in estensione. Il software Real-Time Analysis (RTA) esegue l'analisi delle immagini e l'identificazione delle basi e assegna un punteggio qualitativo a ciascuna base per ogni ciclo di sequenziamento. Al termine dell'analisi primaria, l'analisi secondaria dello strumento MiSeqDx elabora le identificazioni delle basi.

L'elaborazione include solitamente demultiplex, generazione di file FASTQ, allineamento, identificazione di varianti e creazione di file nel formato per l'identificazione delle varianti (VCF) che contengono le informazioni relative alle varianti individuate in specifiche posizioni in un genoma di riferimento. Lo strumento MiSeqDx utilizza moduli diversi per l'analisi secondaria, in base al flusso di lavoro.

## Configurazione dual boot

La configurazione "dual boot" include l'hardware, il software e le procedure di installazione per consentire allo strumento MiSeqDx di eseguire saggi di sequenziamento sia a fini di diagnostica *in vitro* (IVD) sia a soli fini di ricerca (RUO). La configurazione dual boot consente all'utente di spostarsi dalla modalità diagnostica dello strumento alla modalità a solo uso di ricerca dello strumento e viceversa. Le etichette di identificazione a radio frequenza (Radio Frequency Identification, RFID) sui materiali di consumo per il sequenziamento impediscono l'utilizzo dei reagenti di sequenziamento RUO nelle corse diagnostiche.

## Limiti della procedura

- Per uso diagnostico *in vitro*.
- I risultati riportati sull'etichetta sono stati ottenuti con pannelli di saggi rappresentativi utilizzando sangue intero periferico o linee cellulari per le prestazioni relative alle linee germinali e tessuto FFPE o linee cellulari FFPE per le prestazioni relative alle varianti somatiche con i reagenti e i moduli software descritti. I moduli Germline Variant (variante germinale) e Somatic Variant (variante somatica) sono stati sviluppati allo scopo di valutare le prestazioni con saggi rappresentativi. Le caratteristiche delle prestazioni sono fornite solo a scopo illustrativo. I test di convalida illustrati servono esclusivamente a esemplificare le funzionalità generali dello strumento e non stabiliscono le funzionalità dello strumento o la sostenibilità rispetto a qualsiasi specifica dichiarazione. Tutti i test diagnostici sviluppati per l'utilizzo su questo strumento richiedono una convalida completa per tutti gli aspetti delle prestazioni.
- Questo prodotto può fornire solo:
  - Output di sequenziamento  $\geq 5$  Gb a una lunghezza di lettura di 2 x 150 bp
  - Letture che attraversano il filtro  $\geq 15$  milioni a una lunghezza di lettura di 2 x 150 bp
  - Basi superiori a Q30  $\geq 80\%$  alla lunghezza di lettura di 2 x 150 bp  
L'80% o più di basi possiede un punteggio qualitativo sulla scala Phred superiore a 30, a indicare un'accuratezza di identificazione delle basi superiore al 99,9%.
- Lo strumento MiSeqDx è stato convalidato solo per il sequenziamento di librerie di DNA umano estratto da sangue intero periferico o tessuto FFPE. Con questo strumento non devono essere utilizzate librerie generate da altri tipi di campioni per la diagnostica *in vitro*. Non sono state stabilite le prestazioni di questo strumento per il sequenziamento di acido nucleico microbico o virale ottenuto da campioni clinici.
- Lo strumento MiSeqDx è destinato alla diagnostica *in vitro* con reagenti o saggi IVD registrati ed elencati, autorizzati o approvati. I limiti e le caratteristiche prestazionali dei reagenti descritti in questo foglio illustrativo si basano su saggi e moduli software rappresentativi. Per i saggi IVD, fare riferimento al foglio illustrativo del saggio specifico per l'uso previsto, le varianti identificate e il tipo di campione.
- Il contenuto delle indel (inserzioni, delezioni e una combinazione delle due) di lunghezza superiore a 25 bp non è allineato dal software del saggio. Di conseguenza, le indel di lunghezza superiore a 25 bp non sono rilevabili dal software del saggio.
- Il sistema è stato convalidato per il rilevamento di varianti di singolo nucleotide (Single Nucleotide Variant, SNV), di delezioni fino a 25 bp e di inserzioni fino a 24 bp quando utilizzato con il software dei moduli Germline Variant e Somatic Variant. Per le identificazioni delle varianti somatiche, sono state rilevate delezioni di 25 bp e inserzioni di 18 bp a una frequenza della variante di 0,05.
- Le letture degli ampliconi con contenuto estremo della variante potrebbero non essere allineate dal software del saggio, risultando in regioni riportate come wild type. Tale contenuto estremo include:
  - Letture contenenti più di tre indel.
  - Letture di lunghezza di almeno 30 bp con contenuto SNV superiore al 4% della lunghezza totale dell'amplicone target (incluso le regioni della sonda).

- Letture di lunghezza inferiore a 30 bp con contenuto SNV superiore al 10% della lunghezza totale dell'amplicone (incluso le regioni della sonda).
- Le varianti ampie, incluse le varianti di più nucleotidi (Multi-Nucleotide Variant, MNV) e ampie indel, potrebbero essere riportate nel file di output VCF come varianti separate più piccole.
- Le varianti delle delezioni potrebbero essere filtrate o non individuate quando si estendono su due ampliconi rilevati sulle tile se la lunghezza della delezione è pari o maggiore alla sovrapposizione tra gli ampliconi rilevati sulle tile.
- Il sistema non può rilevare le indel, se si verificano direttamente accanto a un primer e dove non è presente un amplicone sovrapposto. Per le regioni con ampliconi sovrapposti, il saggio non può rilevare le delezioni quando la regione di sovrapposizione è inferiore alla dimensione della delezione da rilevare. Ad esempio, se la regione di sovrapposizione tra due ampliconi adiacenti è di due (2) basi, il saggio non può rilevare alcuna delezione includendo entrambe quelle basi. Può essere rilevata una delezione di una singola base a una qualsiasi di quelle basi.
- Come per qualsiasi flusso di lavoro di preparazione delle librerie basato su ibridazione, i polimorfismi, le mutazioni, le inserzioni o le delezioni latenti nelle regioni che legano gli oligonucleotidi possono incidere sugli alleli sondati e, di conseguenza, sulle identificazioni effettuate durante il sequenziamento. Ad esempio:
  - Una variante nella fase con una variante nella regione del primer potrebbe non essere amplificata fornendo un falso negativo.
  - Le varianti nella regione del primer potrebbero impedire l'amplificazione dell'allele di riferimento fornendo un'identificazione della variante omozigote errata.
  - Le varianti indel nella regione del primer potrebbero causare un'identificazione falsa positiva al termine della lettura adiacente al primer.
- Le indel potrebbero essere filtrate a causa di distorsioni del filamento se si verificano accanto alla fine di una lettura e sono sottoposte a soft-clipping durante l'allineamento.
- Piccole MNV potrebbero non essere convalidate.
- Non sono state convalidate le varianti del numero di copie o le varianti strutturali, come le fusioni o le traslocazioni.
- Limitazioni specifiche per il modulo Germline:
  - Il sistema MiSeqDx, utilizzato con il modulo Germline Variant (variante germinale), è progettato per ottenere risultati qualitativi per l'identificazione di varianti della linea germinale (ovvero omozigoti, eterozigoti, wild type).
  - Quando utilizzato con il modulo Germline Variant, la copertura minima per l'amplicone necessaria per l'accurata identificazione delle varianti è di 150x. Il numero di campioni e il numero totale di basi target incidono sulla copertura. Il contenuto in GC e altro contenuto genomico possono incidere sulla copertura.
  - La variazione del numero di copie può incidere sulla possibilità che una variante venga identificata come omozigote o eterozigote.

- Le varianti in un determinato contesto ripetitivo sono filtrate nei file VCF. Il filtro RMxN per le ripetizioni viene utilizzato per filtrare le varianti se tutta o parte della sequenza della variante è presente ripetutamente nel genoma di riferimento adiacente alla posizione della variante. Per l'identificazione delle varianti delle linee germinali, sono richieste almeno nove ripetizioni nel riferimento affinché una variante venga filtrata e vengono prese in considerazione solo le ripetizioni con una lunghezza fino a 5 bp (R5x9).
- Limitazioni specifiche per il modulo Somatic:
  - Il sistema MiSeqDx, utilizzato con il modulo Somatic Variant (variante somatica), è progettato per ottenere risultati qualitativi per l'identificazione delle varianti somatiche (ovvero presenza di una variante somatica con una frequenza di variante pari o superiore a 0,026 con un limite di rilevamento di 0,05).
  - Quando utilizzato con il modulo Somatic Variant, la copertura minima per l'amplicone necessaria per l'accurata identificazione delle varianti è di 450x per il raggruppamento in pool degli oligonucleotidi. Il numero di campioni e il numero totale di basi target incidono sulla copertura. Il contenuto in GC e altro contenuto genomico possono incidere sulla copertura.
  - Le varianti in un determinato contesto ripetitivo sono filtrate nei file VCF. Il filtro RMxN per le ripetizioni viene utilizzato per filtrare le varianti se tutta o parte della sequenza della variante è presente ripetutamente nel genoma di riferimento adiacente alla posizione della variante. Per l'identificazione delle varianti somatiche, sono richieste almeno sei ripetizioni nel riferimento affinché una variante venga filtrata e vengono prese in considerazione solo le ripetizioni con una lunghezza fino a 3 bp (R3x6).
  - Il modulo Somatic Variant non è in grado di differenziare tra le varianti delle linee germinali e le varianti somatiche. Il modulo è progettato per rilevare le varianti su un intervallo di frequenze delle varianti, ma la frequenza della variante non può essere utilizzata per differenziare le varianti somatiche dalle varianti della linea germinale.
  - I tessuti normali contenuti nel campione incidono sul rilevamento delle varianti. Il limite del rilevamento riportato si basa su una frequenza della variante relativa al DNA totale estratto sia da tessuto di tumore che da tessuto normale.

## Componenti del prodotto

Illumina MiSeqDx è composto da:

Strumento MiSeqDx (N. di catalogo DX-410-1001)

Per il funzionamento dello strumento MiSeqDx e per l'analisi dei dati sono necessari i seguenti componenti software:

Applicazione software	Funzione	Descrizione
MiSeq Operating Software (MOS)	Controlla il funzionamento dello strumento	L'applicazione software MOS gestisce il funzionamento dello strumento durante il sequenziamento e genera le immagini utilizzate dal software Real-Time Analysis (RTA). Per ulteriori informazioni, consultare <i>Guida di consultazione dello strumento MiSeqDx per MOS v4 (documento n. 200010452)</i> .
Real-Time Analysis (RTA)	Esegue l'analisi primaria	L'applicazione software RTA converte le immagini generate da MOS per ciascuna tile per ciclo della corsa di sequenziamento in file di identificazione delle basi che rappresentano gli input per i moduli di analisi del Local Run Manager. L'applicazione software RTA non contiene un'interfaccia utente.
Local Run Manager	Interfaccia per la selezione del modulo	Il software Local Run Manager è una soluzione integrata nello strumento per la gestione degli utenti, l'esecuzione dell'analisi secondaria e il monitoraggio dello stato. Per maggiori informazioni consultare <i>Guida al software Local Run Manager v4 per MiSeqDx (documento n. 200046657)</i> .

## Conservazione e manipolazione

Elemento	Specifica
Temperatura	Trasporto e stoccaggio: Da -10 °C a 40 °C Condizioni di funzionamento: Da 19 °C a 25 °C
Umidità	Trasporto e stoccaggio: Umidità senza condensa Condizioni di funzionamento: umidità relativa 30-75% (senza condensa)

## Apparecchiature e materiali richiesti, non forniti

### Materiali di consumo per il sequenziamento

MiSeqDx Reagent Kit v3 (N. di catalogo 20037124)

MiSeqDx Reagent Kit v3 Micro (N. di catalogo 20063860)

### Materiali di consumo forniti dall'utente

Prima di avviare una corsa di sequenziamento, accertarsi di avere a disposizione i seguenti materiali di consumo a carico dell'utente.

Materiale di consumo	Scopo
Salviettine imbevute di alcol isopropilico al 70% oppure etanolo al 70%	Pulizia del vetro e del piano portacelle della cella a flusso
Panno da laboratorio a bassissimo rilascio di particelle	Pulizia del piano portacelle
Carta pulente per lenti, 10 x 15 cm ca.	Pulizia della cella a flusso
Tween 20	Lavaggio dello strumento
Pinzette di plastica con punta quadrata (opzionali)	Estrazione della cella a flusso dal relativo contenitore
Acqua da laboratorio	Lavaggio dello strumento

## Linee guida per l'acqua da laboratorio

Per eseguire le procedure dello strumento utilizzare sempre acqua da laboratorio o acqua deionizzata. Non usare mai acqua di rubinetto.

Utilizzare solo acqua da laboratorio o gli equivalenti seguenti:

- Acqua deionizzata
- PW1 Illumina
- Acqua pari a 18 Megaohm (MΩ)
- Acqua Milli-Q
- Acqua Super-Q
- Acqua sterile per biologia molecolare

## Avvertenze e precauzioni



### AVVERTENZA

**Questo set di reagenti contiene materiali chimici potenzialmente pericolosi. L'inalazione, l'ingestione, il contatto con la pelle o con gli occhi possono causare lesioni personali. Durante la manipolazione di materiali pericolosi nei reagenti deve esservi una ventilazione appropriata. Indossare l'attrezzatura protettiva, inclusi protezione per gli occhi, guanti e indumento da laboratorio appropriato per evitare i rischi di esposizione. Manipolare i reagenti usati come rifiuti chimici e smaltirli in base alle leggi e alle regolamentazioni pertinenti a livello regionale, nazionale e locale.** Per ulteriori informazioni ambientali, di salute e di sicurezza, consultare le SDS alla pagina Web [support.illumina.com/sds.html](http://support.illumina.com/sds.html).

- Manipolare tutti i campioni di sangue come se fossero contagiosi per il virus dell'immunodeficienza umana (Human Immunodeficiency Virus, HIV), il virus dell'epatite B (Human Hepatitis B Virus, HBV) o altri agenti patogeni a trasmissione ematica (precauzioni universali).
- Il mancato rispetto delle procedure descritte può produrre risultati errati o una riduzione significativa della qualità del campione.
- Adottare le normali precauzioni di laboratorio. Non pipettare con la bocca. Non mangiare, bere o fumare nelle aree designate per il lavoro. Manipolare i campioni e i reagenti del kit indossando guanti e indumenti da laboratorio monouso. Dopo aver maneggiato i campioni e i reagenti del kit lavarsi bene le mani.
- È necessario adottare pratiche di laboratorio e igiene di laboratorio idonee per impedire la contaminazione di reagenti, strumenti e campioni di DNA genomico con i prodotti della PCR. La contaminazione da PCR può produrre risultati inesatti e inaffidabili.
- Al fine di prevenire la contaminazione, accertarsi che le aree di pre-amplificazione e di post-amplificazione siano dotate di apparecchiatura dedicata (ad es., pipette, punte per pipette, blocchi termici, agitatori e centrifughe).
- Ove appropriato, l'accoppiamento indice-campione deve corrispondere esattamente al layout della piastra stampata. Local Run Manager popola automaticamente i primer indice associati ai nomi dei campioni, se inseriti nel modulo. Si consiglia di verificare gli index primer associati con i campioni prima di avviare la corsa di sequenziamento. La mancata corrispondenza tra il campione e il layout della piastra risulterà in campioni non identificati correttamente e report con risultati errati.
- Si consiglia vivamente l'installazione di un software antivirus, fornito dall'utente, per proteggere il computer dai virus. Per le istruzioni relative all'installazione consultare il manuale d'uso.
- Non utilizzare MiSeqDx se un qualsiasi pannello è rimosso. Il funzionamento dello strumento con un qualsiasi pannello rimosso crea esposizione potenziale a tensioni di rete e tensioni c.c.
- Non toccare il piano portacelle nello scomparto della cella a flusso. Il riscaldatore in questo scomparto funziona a una temperatura compresa tra 22 °C e 95 °C e potrebbe provocare bruciature.
- Lo strumento pesa circa 57 kg ed eventuali cadute o errata manipolazione possono causare lesioni gravi.
- Riferire immediatamente qualsiasi incidente serio relativo a questo prodotto a Illumina e alle autorità competenti degli stati membri nei quali l'utente e/o il paziente sono residenti.

## Istruzioni per l'uso

Le seguenti istruzioni per l'uso dello strumento MiSeqDx richiedono i reagenti forniti nel MiSeqDx Reagent Kit v3.

## Creazione di una corsa con Local Run Manager

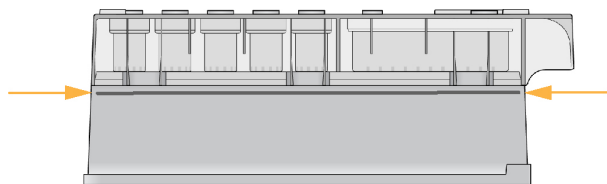
Per istruzioni dettagliate sulla creazione di una corsa, fare riferimento a *Guida al software Local Run Manager v4 per MiSeqDx (documento n. 200046657)* e alla guida per il modulo di Local Run Manager relativa al modulo di analisi utilizzato.

## Preparazione della cartuccia di reagenti

Le istruzioni che seguono descrivono come scongelare i reagenti utilizzando un bagno d'acqua a temperatura ambiente.

1. Estrarre la cartuccia del reagente dal luogo di conservazione a temperatura compresa tra  $-15\text{ }^{\circ}\text{C}$  e  $-25\text{ }^{\circ}\text{C}$ .
2. Collocare la cartuccia di reagenti in un bagno d'acqua contenente abbastanza acqua deionizzata a temperatura ambiente da immergere la base della cartuccia di reagenti fino alla linea di livello acqua stampata sulla cartuccia stessa. Evitare che l'acqua superi la linea di massimo livello acqua.

Figura 1 Linea di massimo livello acqua



3. Lasciare scongelare la cartuccia del reagente nel bagno d'acqua a temperatura ambiente per circa 60-90 minuti o fino a quando è completamente scongelata.
4. Rimuovere la cartuccia dal bagno d'acqua e picchiettarla delicatamente sul banco per far fuoriuscire l'acqua in eccesso dalla base. Asciugare la base della cartuccia. Verificare che sulla parte superiore della cartuccia di reagenti non sia caduta dell'acqua.

## Ispezione della cartuccia di reagenti

1. Capovolgere la cartuccia di reagenti dieci volte per miscelare i reagenti scongelati, quindi ispezionare tutte le posizioni per accertarsi che siano scongelate.
2. Ispezionare i reagenti nelle posizioni 1, 2 e 4 per accertarsi che siano ben miscelati e privi di precipitati.

**NOTA** È fondamentale che i reagenti nella cartuccia siano scongelati completamente e miscelati per assicurare il sequenziamento corretto.

3. Picchiettare delicatamente la cartuccia sul bancone per ridurre le bolle d'aria nei reagenti.

**NOTA** I tubi dei pescanti dello strumento MiSeqDx arrivano fino al fondo di ciascun serbatoio per aspirare i reagenti; per questo motivo è importante che i serbatoi siano privi di bolle d'aria.



4. Riporre la cartuccia in ghiaccio o conservarla a una temperatura compresa tra 2 °C e 8 °C (fino a sei ore) finché non si è pronti a impostare la corsa. Per risultati ottimali, procedere direttamente caricando il campione e impostando la corsa.

## Preparazione dei campioni per il sequenziamento

Per le istruzioni su come preparare le librerie di campioni per il sequenziamento, inclusi la diluizione e il raggruppamento in pool delle librerie, fare riferimento alla sezione Istruzioni per l'uso del foglio illustrativo per la preparazione delle librerie.

La diluizione delle librerie di campioni dipende dalla complessità dei raggruppamenti in pool di oligonucleotidi. Per lo strumento MiSeqDx è necessaria l'ottimizzazione della densità dei cluster; la densità ottimale dei cluster varia a seconda del particolare saggio per la preparazione delle librerie.

## Caricamento delle librerie di campioni sulla cartuccia

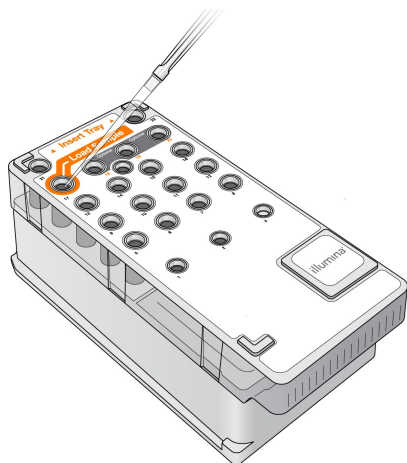
Quando la cartuccia dei reagenti è completamente scongelata e pronta per l'utilizzo, è possibile caricare i campioni sulla cartuccia.

1. Utilizzare una punta per pipette pulita e vuota da 1 ml per forare il sigillo in alluminio sopra il serbatoio sulla cartuccia di reagenti contrassegnato con la dicitura **Load Samples** (Carica campioni).

**NOTA** Non perforare nessun'altra posizione dei reagenti. Le altre posizioni dei reagenti sono perforate automaticamente durante la corsa.

2. Dispensare 600 µl di librerie di campioni DAL (Diluted Amplicon Library) preparate nel serbatoio **Load Samples** (Carica campioni). Evitare di toccare il sigillo in alluminio.
3. Dopo aver caricato il campione, controllare la presenza di bolle d'aria nel serbatoio. In caso di presenza di bolle d'aria, picchiettare delicatamente la cartuccia sul banco in modo da farle fuoriuscire.

Figura 2 Caricamento delle librerie



4. Passare direttamente all'impostazione della corsa usando l'interfaccia MiSeq Operating Software (MOS).

## Impostazione della corsa

Per istruzioni complete sulla configurazione di una corsa fare riferimento a *Guida di consultazione dello strumento MiSeqDx per MOS v4 (documento n. 200010452)*.

1. Accedere a MiSeqDx con la password del software Local Run Manager.
2. Dalla schermata iniziale del software MOS, selezionare **Sequence** (Sequenziamento).
3. Selezionare una corsa dall'elenco, quindi selezionare **Next** (Avanti).  
Una serie di schermate di configurazione della corsa si apre nel seguente ordine: Load Flow Cell (Carica cella a flusso), Load Reagents (Carica reagenti), Review (Revisione) e Pre-Run check (Controllo pre-corsa).
4. Quando viene visualizzata la schermata Load Flow Cell (Carica cella a flusso), pulire e caricare la cella a flusso.
5. Chiudere il coperchio a scatto e lo sportello dello scomparto della cella a flusso.  
Il coperchio a scatto e lo sportello dello scomparto della cella a flusso devono essere chiusi per poter avviare la corsa. Una volta eseguito il caricamento della cella a flusso, il software legge l'etichetta RFID e ne registra il contenuto. Se l'etichetta RFID viene letta correttamente, nell'angolo inferiore destro della schermata appare un messaggio di conferma.
6. Attenersi alle istruzioni del software per caricare il flacone MiSeqDx SBS Solution (PR2), assicurarsi che il flacone degli scarti sia vuoto e caricare la cartuccia di reagenti.  
Quando il flacone MiSeqDx SBS Solution (PR2) e la cartuccia di reagenti sono caricati, il software legge e registra l'etichetta RFID. Se l'etichetta RFID viene letta correttamente, nell'angolo inferiore destro della schermata appare un messaggio di conferma.
7. La schermata Sequencing (Sequenziamento) si apre all'inizio della corsa. La schermata fornisce una rappresentazione visiva dell'avanzamento della corsa, che comprende le intensità e i punteggi qualitativi (Q-scores).

## Risultati

Real-Time Analysis (RTA) è un software integrato che esegue l'analisi delle immagini e l'identificazione delle basi e assegna un punteggio qualitativo a ogni base per ciascun ciclo di sequenziamento. Al termine dell'analisi primaria, il modulo dello strumento MiSeqDx selezionato in [Creazione di una corsa con Local Run Manager alla pagina 8](#) avvia l'analisi secondaria. Per altri flussi di lavoro, fare riferimento alla documentazione specifica del saggio.

## Procedure di controllo qualità

Il software MiSeqDx valuta ciascuna corsa, campione e identificazione delle basi rispetto a metriche di controllo qualità. Se necessario, i controlli positivi e negativi inclusi nella preparazione delle librerie devono essere valutati anche rispetto ai risultati previsti.

## Caratteristiche delle prestazioni

Tutti gli studi sono stati eseguiti con lo strumento MiSeqDx.

Gli studi relativi al modulo Germline Variant hanno utilizzato il saggio MiSeqDx Cystic Fibrosis 139-Variant o i reagenti TruSeq Custom Amplicon Kit Dx per la preparazione delle librerie. I due kit utilizzano reagenti di preparazione delle librerie identici e presentano solo una differenza nel flusso di lavoro: il numero di cicli di reazione a catena della polimerasi (PCR) (25 e 28, rispettivamente). Gli ulteriori cicli di PCR consentono di utilizzare input più bassi di DNA con TruSeq Custom Amplicon Kit Dx (50 ng) per il saggio MiSeqDx Cystic Fibrosis 139-Variant (250 ng), come dimostrato dallo studio sull'input di DNA utilizzando TruSeq Custom Amplicon Kit Dx. Le librerie preparate con il saggio MiSeqDx Cystic Fibrosis 139-Variant sono state sequenziate con i relativi reagenti di sequenziamento contenuti nel kit. Le librerie preparate con TruSeq Custom Amplicon Kit Dx sono state sequenziate con MiSeqDx Reagent Kit v3. I reagenti di sequenziamento di quest'ultimo hanno incrementato gli output rispetto a quelli ottenuti dal saggio MiSeqDx Cystic Fibrosis 139-Variant.

Le analisi comprendono gli intervalli di processività dei campioni supportati da MiSeqDx Reagent Kit v3 Micro. Lo strumento MiSeqDx può supportare da 1 a 96 campioni per corsa, a seconda del saggio. MiSeqDx Reagent Kit v3 Micro è progettato per supportare una processività di campioni inferiore all'interno di questo intervallo per determinati saggi.

Per gli studi somatici è stato utilizzato TruSeq Custom Amplicon Kit Dx con MiSeqDx Reagent Kit v3.

Sia il flusso di lavoro per le linee germinali che il flusso di lavoro per le varianti somatiche, descritti per la preparazione delle librerie per il sequenziamento con TruSeq Custom Amplicon Kit Dx, sono stati seguiti dall'analisi utilizzando il modulo Germline Variant o Somatic Variant, rispettivamente, con due eccezioni. Gli studi che hanno utilizzato un solo gene (prestazioni per le linee germinali; saggio MiSeqDx Cystic Fibrosis 139-Variant) o due geni (prestazioni per le varianti somatiche) come pannelli rappresentativi delle mutazioni hanno utilizzato flussi di lavoro e moduli di analisi specifici per il saggio.

**NOTA** Il contenuto genomico dell'amplicone è riepilogato in relazione al filamento genomico sequenziato. Per gli ampliconi progettati sul filamento minus, il contenuto del genoma di riferimento è il complemento inverso (ad esempio, regioni PolyA su ampliconi del filamento minus corrispondono alle regioni PolyT sul genoma di riferimento).

## Definizione dei calcoli utilizzati nelle caratteristiche delle prestazioni

- La concordanza positiva espressa in percentuale (PPA) è calcolata come la proporzione dei loci classificati come varianti mediante un metodo di riferimento riportati correttamente dal saggio.
  - $(n. \text{ di loci di varianti riportato correttamente dal saggio}) / (n. \text{ totale di loci di varianti})$   
I loci delle varianti riportati dal saggio che concordano con il metodo di riferimento sono veri positivi (True Positive, TP). I loci delle varianti riportati dal saggio come identificazioni di riferimento o come identificazioni delle varianti diverse sono falsi negativi (False Negative, FN).
- La concordanza negativa espressa in percentuale (NPA) è calcolata come la proporzione dei loci classificati come wild type mediante un metodo di riferimento riportati correttamente dal saggio.
  - $(n. \text{ di loci wild type riportato correttamente dal saggio}) / (n. \text{ totale di loci wild type})$   
I loci wild type riportati dal saggio che concordano con il metodo di riferimento sono veri negativi (TN). I loci wild type riportati come varianti dal saggio sono falsi positivi (False Positive, FP).
- La concordanza complessiva espressa in percentuale (OPA) è calcolata come la proporzione dei loci riportati correttamente dal saggio in base a un metodo di riferimento.
  - $[(n. \text{ di loci delle varianti riportato correttamente dal saggio}) + (n. \text{ di loci wild type riportato correttamente dal saggio})] / [(n. \text{ totale di loci delle varianti}) + (n. \text{ totale di loci wild type})]$
- Per le applicazioni di identificazione di varianti, i calcoli di PPA, NPA e OPA non includono le identificazioni non rilevate (i loci delle varianti o dei riferimenti che non soddisfano uno o più filtri di qualità). Due studi hanno specificatamente incluso le identificazioni non riuscite nelle metriche "% di identificazioni corrette" e questa inclusione di identificazioni non riuscite è stata annotata nelle tabelle rilevanti.
- La percentuale di identificazione è calcolata come il numero totale di loci che attraversano i filtri diviso per il numero totale delle posizioni sequenziate per i cromosomi 1-22. I cromosomi X e Y sono esclusi. Questa metrica non prende in considerazione la concordanza delle identificazioni rilevate dal metodo di riferimento.

Per le caratteristiche prestazionali correlate a fattori pre-analitici (es. metodi di estrazione o DNA input), fare riferimento al foglio illustrativo per il metodo di preparazione delle librerie pertinente.

## Indicizzazione dei campioni

Gli index primer dei campioni, aggiunti durante la preparazione delle librerie, assegnano una sequenza univoca a ciascun campione di DNA, consentendo il raggruppamento in pool di più campioni in una singola corsa di sequenziamento. Il test di indicizzazione dei campioni è stato eseguito sia per il flusso di lavoro Germline che per il flusso di lavoro Somatic.

Sono stati analizzati complessivamente 96 indici di campioni mediante un saggio rappresentativo progettato per interrogare diversi geni che coprono 12.588 basi per filamento su tutti i 23 cromosomi umani, al fine di verificare la capacità del saggio di identificare i genotipi in modo coerente per un dato campione su diverse combinazioni di primer di indicizzazione. Il cromosoma Y non contiene regioni affidabili e non è stato valutato.

Sono stati analizzati otto campioni univoci con 12 diverse combinazioni di primer di indicizzazione per campione. I risultati ottenuti da modulo di analisi Germline sono stati confrontati con Platinum Genomes versione 2016-01. La PPA (SNV e indel) ha superato il 97% (le identificazioni vero positive erano almeno 70 per le SNV, 38 per le inserzioni, 36 per le delezioni) e la NPA era del 100% (almeno 23.440 posizioni di riferimento per combinazione indice) per ciascuna delle 96 combinazioni indice. Indipendentemente è stato analizzato un singolo indice per verificare che la chimica di sequenziamento di MiSeqDx Reagent Kit v3 è in grado di supportare meno di otto campioni (la chimica del precedente MiSeqDx Universal Kit 1.0 si limitava a un minimo di otto campioni). Il singolo indice presentava valori di PPA del 98,9% (180/182) per le SNV, 100% (38/38) per le inserzioni e 100% (46/46) per le delezioni. La NPA era del 100% (23.856/23.856).

Sono stati analizzati dodici replicati (24 librerie) di un campione per misurare l'accuratezza dell'indice con le varianti somatiche a un intervallo di frequenze di 0,05-0,10 utilizzando il modulo Somatic Variant (per identificare le varianti somatiche sono state utilizzate due combinazioni indici per replicato). La PPA era del 100% per le SNV (64/64), le inserzioni (11/11) e le delezioni (19/19). La NPA era del 100% (almeno 11.590 posizioni di riferimento per combinazione indici) per tutte le combinazioni indici.

## Carry-over dei campioni

Il flusso di lavoro dello strumento MiSeqDx prevede la preparazione delle librerie e il sequenziamento simultaneo di molteplici campioni e controlli. È stato condotto uno studio del carry-over dei campioni per valutare se i risultati falsi positivi, dovuti al carry-over causato dalla contaminazione tra pozzetti durante la preparazione delle librerie di campioni nonché dalla contaminazione tra corse in caso di corse di sequenziamento consecutive, incidano sui risultati dell'analisi. Sono state utilizzate le varianti somatiche, in quanto possono essere rilevate in eventi di frequenze alleliche inferiori rispetto alle varianti delle linee germinali.

Sono stati utilizzati quattro campioni di DNA genomico ottenuto da linee cellulari, ciascuno contenente mutazioni diverse nel pannello in un saggio rappresentativo di due geni. I campioni erano tali che una mutazione in una posizione del campione presenterà una sequenza di riferimento (wild type) nelle altre.

Il carry-over da pozzetto a pozzetto è definito come una procedura non riuscita potenzialmente creata dall'elaborazione manuale (pipettamento, mescolamento dei campioni e così via). Per valutare il carry-over da un pozzetto di campione all'altro, sono state eseguite due corse di analisi:

- Un layout a scacchiera di un campione a elevato input di DNA genomico (gDNA) contenente una mutazione nel Gene 1 alternato con un campione a basso input di gDNA contenente una mutazione nel Gene 2.
- Un layout a scacchiera con un campione di gDNA input alto contenente una mutazione nel Gene 2 alternato a un campione di gDNA input basso contenente una mutazione nel Gene 1.

In ciascuna corsa è stato valutato un totale di 12 replicati per individuare la presenza di falsi positivi (ad es., una mutazione nel Gene 1 riportata in un pozzetto designato come un campione contenente la mutazione nel Gene 2 o vice versa).

Il carry-over da corsa a corsa è definito come una procedura non riuscita potenzialmente creata da residuo appartenente a una precedente corsa di sequenziamento. Per determinare l'esistenza di carry-over tra le corse di sequenziamento sono state preparate e sequenziate consecutivamente su uno strumento MiSeqDx due

piastre, ciascuna contenente 11 replicati di un singolo campione univoco di gDNA input alto più un campione bianco, ed è stata valutata la presenza di falsi positivi. La prima corsa conteneva 11 replicati di un campione contenente la mutazione nel Gene 2 più un campione bianco. La seconda corsa conteneva 11 replicati di un campione contenente la mutazione nel Gene 1 più un campione bianco. La libreria del campione contenente la mutazione del Gene 2 è stata sequenziata per prima seguita da una successiva corsa di sequenziamento con la libreria del campione contenente la mutazione del Gene 1, seguita da un'altra ripetizione della corsa di sequenziamento della libreria del campione contenente la mutazione del Gene 2. Se viene osservata una qualsiasi mutazione del Gene 2 in una corsa con la sola mutazione del Gene 1, e vice versa, questa osservazione indicherebbe la presenza di carry-over.

Non è stato riportato alcun falso positivo (0/24, 0%) a causa di carry-over *da pozzetto a pozzetto*. Sono state rilevate tutte le mutazioni previste. Non è stato riportato alcun falso positivo (0/24, 0%) a causa di carry-over *da corsa a corsa*. Sono state rilevate tutte le mutazioni previste. Non è stato riportato alcun falso positivo (0/48, 0%) a causa di carry-over *totale* (carry-over combinato da pozzetto a pozzetto e da corsa a corsa).

## Caratteristiche delle prestazioni per le linee germinali

Gli studi qui descritti hanno utilizzato il modulo Germline Variant per analizzare i dati del sequenziamento, fatta eccezione per quegli studi che hanno utilizzato un pannello di singolo gene dove è stato utilizzato un modulo specifico per il saggio.

### Accuratezza

È stato condotto lo studio seguente per valutare l'accuratezza dello strumento MiSeqDx con MiSeqDx Reagent Kit v3 e DNA di alta qualità. Lo studio ha utilizzato un saggio rappresentativo progettato per interrogare diversi geni che coprono 12.588 basi su 23 cromosomi diversi utilizzando 150 ampliconi. Il cromosoma Y non contiene regioni affidabili e non è stato valutato. I dodici campioni univoci utilizzati in questo studio appartengono a una singola famiglia (due genitori e dieci bambini). Questi campioni sono stati spesso sequenziati da più laboratori e con diverse metodologie di sequenziamento. Cinque campioni sono femminili e sette sono maschili. Ciascun campione è stato analizzato in duplicati. L'accuratezza è stata determinata per SNV, inserzioni e delezioni confrontando i dati dello studio su un database di riferimenti ben caratterizzati. Le sequenze di riferimento del database (Platinum Genomes versione 2016-01) sono state derivate dalla combinazione di più metodologie di sequenziamento, dati disponibili pubblicamente e informazioni ereditarie. Le regioni genomiche affidabili sono state definite in base a questo metodo di riferimento, se non altrimenti specificato. In tutto i campioni sono stati analizzati otto volte; le tabelle dimostrano che l'accuratezza si basava sui dati ottenuti dalla prima corsa.

La [Tabella 1](#) contiene i dati dello studio con la concordanza positiva e negativa espressa in percentuale in base ai singoli campioni, dove i risultati delle varianti sono confrontati con un metodo di riferimento composito ben caratterizzato per i calcoli della PPA. Sono stati combinati i tre tipi di varianti (SNV, inserzioni e delezioni). Poiché il metodo di riferimento fornisce solo risultati per le varianti di singolo nucleotide e inserzioni/delezioni, i risultati delle basi non varianti sono confrontati con la sequenza di riferimento del genoma umano versione hg19, per i calcoli della NPA.

Tabella 1 Concordanza dei risultati di identificazione delle basi dello strumento MiSeqDx per campione

Campione	Percentuale di identificazione media	N. totale di varianti	N. totale di varianti TP	N. totale di varianti FN	N. totale di identificazioni non riuscite	N. totale di identificazioni TN	PPA	NPA	OPA
NA12877	>99,9%	152	152	0	4	24.024	100	100	100

Campione	Percentuale di identificazione media	N. totale di varianti	N. totale di varianti TP	N. totale di varianti FN	N. totale di identificazioni non riuscite	N. totale di identificazioni TN	PPA	NPA	OPA
NA12878	>99,9%	270	266	0	4	23.856	100	100	100
NA12879	>99,9%	192	190	1	1	24.054	99,5	100	>99,9%
NA12880	>99,9%	222	220	0	6	24.052	100	100	100
NA12881	>99,9%	250	247	1	2	23.862	99,6	100	>99,9%
NA12882	>99,9%	200	196	2	2	23.962	99,0	100	>99,9%
NA12883	>99,9%	226	224	0	6	23.870	100	100	100
NA12884	>99,9%	228	226	1	1	23.942	99,6	100	>99,9%
NA12885	>99,9%	244	240	2	2	23.942	99,2	100	>99,9%
NA12886	>99,9%	230	228	1	1	23.888	99,6	100	>99,9%
NA12888	>99,9%	216	216	0	4	24.002	100	100	100
NA12893	>99,9%	236	234	0	2	23.810	100	100	100

Il saggio rappresentativo consisteva di 150 ampliconi progettati per coprire diverso contenuto genomico. Il contenuto in GC degli ampliconi rientrava nell'intervallo compreso tra il 26% e l'87%. Gli ampliconi presentavano inoltre un intervallo di ripetizioni di singolo nucleotide (ad es., PolyA, PolyT), dinucleotide e trinucleotide. I dati sono stati compilati in base ai singoli ampliconi (Tabella 2), allo scopo di determinare gli effetti del contenuto genomico sulla percentuale di identificazioni corrette. La percentuale di identificazioni corrette consiste nell'identificazione di varianti e riferimenti ed è inferiore al 100% in presenza di identificazioni errate o non riuscite. L'identificazione non riuscita si verifica quando l'identificazione di varianti non soddisfa uno o più filtri (ad es., copertura insufficiente).

Delle otto varianti FN della Tabella 2, sette corrispondevano a un'inserzione di 1 bp nell'amplicone 111 che contiene anche l'omopolimero PolyA e presenta un contenuto in GC di 0,29. La variante FN restante (chiamata errata) era dovuta a una SNV eterozigote attesa sull'amplicone 125 con un contenuto GC di 0,68, identificata come variante omozigote. La frequenza della variante SNV era di 0,71, al di sopra della soglia di 0,70 per la classificazione come variante omozigote. L'amplicone con la percentuale inferiore di identificazioni corrette (98,2%) era l'amplicone 17 con 40 identificazioni non riuscite e contenente le ripetizioni AT e il contenuto in GC di 27%.



Tabella 2 Accuratezza a livello di amplicone per lo strumento MiSeqDx

Amplicone	Cromosoma	Avvio amplicone	Fine amplicone	Dimensione del frammento analizzato	Basi nella regione affidabile	Contenuto genomico dell'amplicone	Contenuto in GC	N. totale di identificazioni corrette	N. totale di identificazioni errate	N. totale di identificazioni non riuscite	% di identificazioni corrette
1	1	36450499	36450591	93	93	Indel	0,22	2.232	0	0	100
2	1	109465122	109465200	79	79	PolyA (5), PolyC (5), indel	0,38	1.896	0	0	100
3	1	218353867	218353957	91	91	Indel	0,4	2.184	0	0	100
4	1	223906657	223906748	92	92	Indel	0,49	2.208	0	0	100
5	1	228526602	228526682	81	81	PolyG (5)	0,69	1.944	0	0	100
6	1	236372039	236372108	70	70	PolyT (10), indel	0,39	1.680	0	0	100
7	1	247812041	247812128	88	88	PolyA (5), CT(3), TAA(3), indel	0,27	2.112	0	0	100
8	2	55862774	55862863	90	90	Indel	0,28	2.160	0	0	100
9	2	87003930	87004009	80	80	Indel	0,38	1.920	0	0	100
10	2	177016721	177016805	85	81	N/D	0,65	1.944	0	0	100
11	2	186625727	186625801	75	75	PolyA (8)	0,35	1.800	0	0	100
12	2	190323504	190323591	88	88	PolyT (5)	0,42	2.112	0	0	100
13	2	200796740	200796826	87	87	PolyT (5), indel	0,31	2.088	0	0	100
14	2	212245049	212245139	91	91	PolyT (5), PolyA (6), indel	0,3	2.184	0	0	100
15	2	228147052	228147144	93	93	N/D	0,43	2.232	0	0	100
16	2	235016350	235016422	73	73	PolyT (5), indel	0,42	1.752	0	0	100
17	3	4466229	4466321	93	93	AT(3), indel	0,27	2.192	0	40	98,2
18	3	46620561	46620643	83	83	N/D	0,43	1.992	0	0	100
19	3	49851331	49851400	70	70	CT(3), indel	0,49	1.680	0	0	100
20	3	189713161	189713248	88	88	PolyA (5), PolyT (5), PolyA (9), TG (3)	0,41	2.112	0	0	100
21	3	190106030	190106104	75	74	Indel	0,57	1.774	0	2	99,9
22	4	2233667	2233744	78	78	PolyA (6)	0,26	1.872	0	0	100
23	4	7780541	7780637	97	97	PolyG (6), PolyT (5), PolyA (5)	0,42	2.328	0	0	100

Amplicone	Cromosoma	Avvio amplicone	Fine amplicone	Dimensione del frammento analizzato	Basi nella regione affidabile	Contenuto genomico dell'amplicone	Contenuto in GC	N. totale di identificazioni corrette	N. totale di identificazioni errate	N. totale di identificazioni non riuscite	% di identificazioni corrette
24	4	15688604	15688681	78	78	N/D	0,29	1.872	0	0	100
25	4	56236521	56236586	66	62	PolyA (5), indel	0,36	1.488	0	0	100
26	4	102839244	102839314	71	69	PolyA (5)	0,46	1.656	0	0	100
27	4	164446743	164446804	62	62	PolyA (7), indel	0,27	1.488	0	0	100
28	5	1882081	1882158	78	75	N/D	0,78	1.800	0	0	100
29	5	14769061	14769144	84	84	GT(3), CCA(3)	0,62	2016	0	0	100
30	5	41069808	41069871	64	64	N/D	0,39	1.536	0	0	100
31	5	74077114	74077196	83	83	PolyA (6), indel	0,3	1.992	0	0	100
32	5	147475343	147475409	67	67	PolyT (5)	0,37	1.608	0	0	100
33	5	149323731	149323821	91	91	CT(4), AG(3)	0,55	2.184	0	0	100
34	5	155662213	155662287	75	75	Indel	0,43	1.800	0	0	100
35	6	6318713	6318814	10	10	PolyG (6)	0,68	2.448	0	0	100
36	6	24949983	24950074	92	92	Indel	0,63	2.208	0	0	100
37	6	31084900	31084999	100	94	GCT(5), indel	0,61	2.244	0	12	99,5
38	6	32147987	32148084	98	98	PolyT (5), TCT(3), CTT(3)	0,55	2.352	0	0	100
39	6	32986864	32986958	95	95	Indel	0,53	2.280	0	0	100
40	6	33408498	33408583	86	86	PolyC (6)	0,7	2064	0	0	100
41	6	41647401	41647495	95	94	PolyG (5), indel	0,61	2.256	0	0	100
42	6	112435865	112435955	91	91	PolyA (5)	0,44	2.184	0	0	100
43	7	22202076	22202148	73	73	N/D	0,44	1.752	0	0	100
44	7	66276100	66276187	88	88	Indel	0,35	2.112	0	0	100
45	7	77365735	77365821	87	87	PolyA (7), AG(4)	0,26	2.088	0	0	100
46	7	110939946	110940030	85	85	Indel	0,38	2.040	0	0	100
47	7	128533468	128533557	90	90	PolyG (5), indel	0,62	2.160	0	0	100
48	7	149503875	149503965	91	91	PolyG (6), PolyC (6), indel	0,71	2.184	0	0	100
49	7	154404519	154404599	81	66	N/D	0,31	1.584	0	0	100
50	7	156476507	156476599	93	93	Indel	0,35	2.232	0	0	100

Amplicone	Cromosoma	Avvio amplicone	Fine amplicone	Dimensione del frammento analizzato	Basi nella regione affidabile	Contenuto genomico dell'amplicone	Contenuto in GC	N. totale di identificazioni corrette	N. totale di identificazioni errate	N. totale di identificazioni non riuscite	% di identificazioni corrette
51	8	1817312	1817394	83	83	N/D	0,42	1.992	0	0	100
52	8	24811020	24811109	90	89	PolyG (7), CTC(4), indel	0,61	2113	0	23	98,9
53	8	76518625	76518691	67	67	Indel	0,3	1.608	0	0	100
54	9	103054909	103055006	98	98	PolyG (6)	0,67	2.352	0	0	100
55	9	105586150	105586214	65	65	Indel	0,32	1.560	0	0	100
56	9	107620823	107620918	96	96	N/D	0,49	2.304	0	0	100
57	9	123769149	123769231	83	83	AT(3)	0,37	1.992	0	0	100
58	9	138995345	138995441	97	97	PolyC (6), indel	0,68	2.328	0	0	100
59	10	5987120	5987198	79	78	PolyG (5), indel	0,47	1.872	0	0	100
60	10	11784629	11784726	98	91	GC(3)	0,87	2.184	0	0	100
61	10	27317777	27317855	79	79	PolyT (5)	0,3	1.896	0	0	100
62	10	33018351	33018440	90	90	PolyA (5), PolyT (5)	0,2	2.160	0	0	100
63	10	45084159	45084253	95	95	Indel	0,35	2.280	0	0	100
64	10	55892599	55892687	89	88	AC(11), indel	0,42	2.102	0	10	99,5
65	10	101611250	101611329	80	80	N/D	0,49	1.920	0	0	100
66	10	118351373	118351453	81	81	N/D	0,51	1.944	0	0	100
67	11	8159816	8159912	97	96	N/D	0,45	2.304	0	0	100
68	11	30177648	30177717	70	70	Indel	0,46	1.680	0	0	100
69	11	47470345	47470444	100	100	N/D	0,65	2.400	0	0	100
70	11	59837679	59837740	62	62	Indel	0,37	1.488	0	0	100
71	11	64418856	64418957	102	102	N/D	0,59	2.448	0	0	100
72	11	93529612	93529684	73	73	PolyA (5)	0,4	1.752	0	0	100
73	11	101347052	101347136	85	85	N/D	0,42	2.040	0	0	100
74	11	102477336	102477426	91	91	PolyG (6)	0,55	2.184	0	0	100
75	11	118406285	118406369	85	85	Indel	0,53	2.040	0	0	100
76	11	120357801	120357885	85	85	PolyA (5), CA(3), indel	0,34	2.040	0	0	100

Amplicone	Cromosoma	Avvio amplicone	Fine amplicone	Dimensione del frammento analizzato	Basi nella regione affidabile	Contenuto genomico dell'amplicone	Contenuto in GC	N. totale di identificazioni corrette	N. totale di identificazioni errate	N. totale di identificazioni non riuscite	% di identificazioni corrette
77	11	125769313	125769397	85	85	GA(3)	0,52	2.040	0	0	100
78	12	2834770	2834853	84	84	PolyC (5), indel	0,52	2016	0	0	100
79	12	26811004	26811096	93	93	PolyA (7), AC(4)	0,33	2.232	0	0	100
80	12	30881766	30881846	81	81	N/D	0,49	1.944	0	0	100
81	12	88474105	88474175	71	71	PolyA (6)	0,35	1.704	0	0	100
82	12	120966872	120966966	95	95	PolyG (5)	0,68	2.280	0	0	100
83	13	24167504	24167576	73	73	N/D	0,52	1.752	0	0	100
84	13	25816961	25817049	89	88	PolyA (5), PolyT (7), PolyA (7), indel	0,22	2.112	0	0	100
85	13	44880112	44880200	89	89	Indel	0,49	2.136	0	0	100
86	13	77665218	77665294	77	77	Indel	0,39	1.848	0	0	100
87	14	31619327	31619393	67	67	GA(3), TA(3)	0,39	1.608	0	0	100
88	14	39517884	39517966	83	83	N/D	0,25	1.992	0	0	100
89	14	46958962	46959034	73	72	PolyT (5), indel	0,19	1.727	0	1	99,9
90	14	58050030	58050110	81	81	Indel	0,38	1.944	0	0	100
91	14	82390559	82390649	91	91	Indel	0,35	2.184	0	0	100
92	14	92549544	92549609	66	66	PolyA (5)	0,41	1.584	0	0	100
93	14	102808496	102808589	94	94	Indel	0,62	2.256	0	0	100
94	15	43170751	43170848	98	96	PolyC (5)	0,45	2.304	0	0	100
95	15	63446149	63446216	68	68	Indel	0,25	1.632	0	0	100
96	15	77879807	77879901	95	93	PolyG (5), indel	0,68	2.232	0	0	100
97	15	81625334	81625428	95	95	PolyT (6)	0,43	2.280	0	0	100
98	15	85438263	85438334	72	71	Indel	0,65	1.704	0	0	100
99	15	89817413	89817503	91	91	N/D	0,36	2.184	0	0	100
100	15	89864274	89864343	70	70	Indel	0,56	1.680	0	0	100
101	16	1894910	1894972	63	63	N/D	0,27	1.512	0	0	100
102	16	28997904	28997998	95	95	PolyC (5)	0,67	2.280	0	0	100
103	16	53682908	53682994	87	87	TA(3)	0,41	2.088	0	0	100
104	16	57954406	57954509	104	104	PolyC (5)	0,67	2.496	0	0	100

Amplicone	Cromosoma	Avvio amplicone	Fine amplicone	Dimensione del frammento analizzato	Basi nella regione affidabile	Contenuto genomico dell'amplicone	Contenuto in GC	N. totale di identificazioni corrette	N. totale di identificazioni errate	N. totale di identificazioni non riuscite	% di identificazioni corrette
105	16	85706375	85706465	91	91	Poly T (5), indel	0,37	2.184	0	0	100
106	17	3563920	3564008	89	89	GC(3)	0,64	2.136	0	0	100
107	17	3594191	3594277	87	87	PolyC (5), indel	0,67	2.088	0	0	100
108	17	3970090	3970180	91	91	Indel	0,46	2.184	0	0	100
109	17	16084945	16085037	93	93	Indel	0,26	2.232	0	0	100
110	17	33998759	33998849	91	89	PolyT (5)	0,54	2.136	0	0	100
111	17	39589691	39589774	84	82	PolyA (13), indel (x2)	0,29	1.944	7	17	98,8
112	17	41244394	41244484	91	91	PolyA (5)	0,34	2.184	0	0	100
113	17	45438866	45438957	92	92	PolyA (7), AT(3), AT(4), AT(4), indel	0,26	2.208	0	0	100
114	17	61502432	61502510	79	79	Indel	0,41	1.887	0	9	99,5
115	17	64023582	64023667	86	86	PolyT (7)	0,22	2064	0	0	100
116	17	72308237	72308320	84	84	GAG(3)	0,62	2016	0	0	100
117	18	2616456	2616522	67	67	GA(3)	0,31	1.608	0	0	100
118	18	6980478	6980568	91	91	N/D	0,37	2.184	0	0	100
119	18	9888026	9888094	69	69	PolyA (6), TG(3)	0,43	1.656	0	0	100
120	18	38836999	38837073	75	75	PolyA (5), indel	0,37	1.800	0	0	100
121	18	47405382	47405462	81	81	CTC(3), indel	0,47	1.944	0	0	100
122	18	54815665	54815749	85	85	CT(3), indel	0,45	2.040	0	0	100
123	18	59773996	59774060	65	65	N/D	0,48	1.560	0	0	100
124	19	625143	625241	99	99	N/D	0,59	2.376	0	0	100
125	19	18121418	18121491	74	74	N/D	0,68	1.775	1	0	99,9
126	19	18186574	18186643	70	70	N/D	0,64	1.680	0	0	100
127	20	746056	746149	94	94	N/D	0,61	2.256	0	0	100
128	20	10633195	10633276	82	82	AC(3)	0,59	1.968	0	0	100
129	20	17705633	17705708	76	76	CT(3)	0,58	1.824	0	0	100
130	20	21766821	21766890	70	70	GT(3), TG(4), indel	0,46	1.680	0	0	100
131	20	25278421	25278521	101	101	Indel	0,63	2.424	0	0	100

Amplicone	Cromosoma	Avvio amplicone	Fine amplicone	Dimensione del frammento analizzato	Basi nella regione affidabile	Contenuto genomico dell'amplicone	Contenuto in GC	N. totale di identificazioni corrette	N. totale di identificazioni errate	N. totale di identificazioni non riuscite	% di identificazioni corrette
132	20	50897302	50897368	67	67	Indel	0,36	1.608	0	0	100
133	20	62331904	62331994	91	88	PolyG (6)	0,73	2.112	0	0	100
134	20	62690860	62690946	87	87	Indel	0,57	2.088	0	0	100
135	21	30300823	30300888	66	66	Indel	0,35	1.584	0	0	100
136	21	33694176	33694273	98	98	PolyT (6), CA(3)	0,54	2.352	0	0	100
137	21	36710706	36710792	87	87	GT(3), indel	0,39	2.088	0	0	100
138	21	46644924	46644992	69	69	PolyA (6), AG(3), indel	0,32	1.656	0	0	100
139	21	46705575	46705664	90	90	PolyT (5), PolyA (6)	0,5	2.160	0	0	100
140	22	25750774	25750873	100	100	Indel	0,63	2.400	0	0	100
141	22	32439233	32439329	97	97	N/D	0,68	2.328	0	0	100
142	22	37409844	37409940	97	97	Indel	0,46	2.328	0	0	100
143	22	37637596	37637694	99	99	N/D	0,6	2.376	0	0	100
144	22	47081347	47081438	92	92	Indel	0,66	2.208	0	0	100
145	X	15870424	15870492	69	69	PolyT (5)	0,26	1.656	0	0	100
146	X	135288543	135288611	69	69	PolyC (5)	0,62	1.656	0	0	100
147	X	135290777	135290847	71	71	N/D	0,52	1.704	0	0	100
148	Y	2655397	2655461	65	0	N/D	0,55	0	0	0	N/D
149	Y	2655519	2655609	91	0	N/D	0,48	0	0	0	N/D
150	Y	2655609	2655679	71	0	PolyA (5)	0,37	0	0	0	N/D

Le varianti con identificazione non riuscita sono riepilogate nella [Tabella 3](#). I filtri che hanno fornito le identificazioni non riuscite sono elencati nella tabella. L'inserzione sull'amplicone 111 è stata filtrata per nove delle 16 occorrenze con le restanti sette occorrenze identificate come riferimento e sono quindi FN.

Tabella 3 Riepilogo delle varianti con identificazione non riuscita

N. amplicone	Crom:Pos	Variante	Contenuto dell'amplicone corrispondente	Filtro	Varianti non identificate	Varianti previste	Identificazioni FN
64	10:55892600	TAC > T	AC(11), 42% GC	R5x9 <sup>1</sup>	10	10	0
111	17:39589692	C > CA	PolyA (13), 29% GC	R5x9	9	16	7

<sup>1</sup> R5x9: ripetere il filtro. Una variante viene filtrata se tutta o parte della variante è presente ripetutamente nel genoma di riferimento adiacente alla posizione della variante. Sono richieste almeno nove ripetizioni nel riferimento e sono prese in considerazione solo le ripetizioni con lunghezza fino a 5 bp.

I risultati del sequenziamento per il campione NA12878 sono stati confrontati con un genotipo altamente affidabile per NA12878, come stabilito dal National Institutes of Standards and Technology (NIST) (v.2.19). Dei 150 ampliconi, 92 ampliconi sono rientrati completamente nelle regioni genomiche altamente affidabili, 41 ampliconi presentavano una sovrapposizione parziale e 17 ampliconi non presentavano alcuna sovrapposizione nella sequenza NIST. Questo ha fornito 10.000 coordinate per replicato per il confronto. Le identificazioni delle basi non varianti sono state confrontate con la sequenza di riferimento del genoma umano versione 19. I risultati dell'accuratezza sono mostrati nella [Tabella 4](#).

Tabella 4 Concordezza dei risultati di identificazione delle basi dello strumento MiSeqDx per il campione NA12878 con il database NIST

Campione	N. di ampliconi	Percentuale di identificazione media	N. totale di varianti TP	N. totale di varianti FN	N. totale di identificazioni TN	N. totale di identificazioni FP	PPA	NPA	OPA
NA12878	133	99,98	208	0	19.380	0	100	100	100

I campioni sono stati ulteriormente analizzati per identificare piccole inserzioni e delezioni (indel) ([Tabella 5](#)). In alcuni casi, l'indel era comune tra due o più campioni, come illustrato nella colonna "N. totale di replicati del campione con indel". I risultati per entrambi i replicati dei 12 campioni validi sono inclusi nella [Tabella 5](#). Era presente un totale di 71 indel con una dimensione compresa nell'intervallo 1-24 bp per le inserzioni e 1-25 bp per le delezioni. Per 69 indel è stata rilevata una percentuale di concordanza positiva del 100%. Una delezione (amplicone 64; delezione di 2 bp (chr10 55892600 TAC>T) non è stata identificata correttamente perché ciascuna di queste

varianti risultava un'identificazione non riuscita a causa del filtro R5x9. Pertanto, non è stato possibile calcolare il PPA che esclude le identificazioni non riuscite. Anche un altro indel, l'inserimento di 1 bp (chr17 39589692 C>CA nell'amplicone 111), non è stata identificata correttamente, perché nove varianti sono risultate un'identificazione non riuscita a causa del filtro R5x9 e sette sono risultate identificazioni FN.

Tabella 5 Riepilogo dell'identificazione di indel con lo strumento MiSeqDx

Amplicone	Cromosoma	Posizione	Dimensione del frammento analizzato	Tipo e lunghezza dell'indel per l'amplicone	Indel	N. totale di replicati del campione con indel	N. di identificazioni non riuscite	N. totale di identificazioni indel errate	N. totale di identificazioni indel corrette	PPA
1	1	36450544	93	Delezione di 25 bp	GAAAATTTAATGAAACACATTGCCT>G	2	0	0	2	100
2	1	109465165	79	Delezione di 3 bp	ACTT>A	12	0	0	12	100
3	1	218353908	91	Inserzione di 23 bp	T>TTTTAATAGCAAAAAGAGGCTAGA	24	0	0	24	100
4	1	223906701	92	Delezione di 17 bp	GACAGACTGTGAGGAAGA>G	10	0	0	10	100
6	1	236372081	70	Inserzione di 5 bp	C>CTTAAG	10	0	0	10	100
7	1	247812083	88	Inserzione di 3 bp	C>CATG	10	0	0	10	100
8	2	55862804	90	Inserzione di 7 bp	T>TTTGTTAA	14	0	0	14	100
9	2	87003972	80	Delezione di 6 bp	TTATCTC>T	6	0	0	6	100
13	2	200796749	87	Inserzione di 5 bp	T>TTAAAA	24	0	0	24	100
14	2	212245090	91	Inserzione di 12 bp	C>CTGAAAAATAGGAT	14	0	0	14	100
16	2	235016388	73	Inserzione di 2 bp	A>ATG	12	0	0	12	100
17	3	4466274	93	Delezione di 23 bp	TAACCTAAAATTACAAAATAACCC>T	2	0	0	2	100
19	3	49851375	70	Inserzione di 9 bp	C>CCTGGCTCCT	4	0	0	4	100
21	3	190106071	75	Delezione di 1 bp	AG>A	20	0	0	20	100
25	4	56236567	66	Delezione di 8 bp	TAACCGAAA>T	12	0	0	12	100
27	4	164446785	62	Inserzione di 11 bp	T>TTATGGTATTGA	12	0	0	12	100
31	5	74077155	83	Delezione di 4 bp	TAGTA>T	10	0	0	10	100
34	5	155662255	75	Inserzione di 8 bp	G>GCCTACTGA	20	0	0	20	100
36	6	24950035	92	Delezione di 21 bp	CCCTGGGTGCTATAGCCACCA>C	10	0	0	10	100
37	6	31084942	100	Delezione di 3 bp	GCTT>G	14	0	0	14	100
39	6	32986905	95	Delezione di 25 bp	CTTTCACTTTCCCGTCTCATGCAAG>C	12	0	0	12	100
41	6	41647442	95	Delezione di 23 bp	GGCATGAGGCTTGGTGACATGGCA>G	8	0	0	8	100



Amplicone	Cromosoma	Posizione	Dimensione del frammento analizzato	Tipo e lunghezza dell'indel per l'amplicone	Indel	N. totale di replicati del campione con indel	N. di identificazioni non riuscite	N. totale di identificazioni indel errate	N. totale di identificazioni indel corrette	PPA
44	7	66276142	88	Inserzione di 1 bp	C>CT	16	0	0	16	100
46	7	110939983	85	Delezione di 4 bp	CAAGT>C	12	0	0	12	100
47	7	128533514	90	Inserzione di 1 bp	T>TC	24	0	0	24	100
48	7	149503916	91	Delezione di 4 bp	GGATA>G	8	0	0	8	100
50	7	156476548	93	Delezione di 11 bp	GAATCTGCACTT>G	12	0	0	12	100
52	8	24811064	90	Delezione di 1 bp	AG>A	24	0	0	24	100
53	8	76518677	67	Inserzione di 4 bp	T>TACTG	14	0	0	14	100
55	9	105586193	65	Inserzione di 4 bp	C>CAATT	2	0	0	2	100
58	9	138995370	97	Delezione di 21 bp	TCTGGGGGGCAGCCCCTGAGGG>T	14	0	0	14	100
59	10	5987158	79	Delezione di 3 bp	TAAC>T	10	0	0	10	100
63	10	45084202	95	Delezione di 16 bp	AGCGTCTATAACCAAAT>A	12	0	0	12	100
64	10	55892600	89	Delezione di 2 bp	TAC>T	10	10	0	0	N/D
68	11	30177690	70	Inserzione di 2 bp	C>CTG	10	0	0	10	100
70	11	59837721	62	Inserzione di 8 bp	T>TTATGAAAA	12	0	0	12	100
75	11	118406328	85	Delezione di 8 bp	CAGTGTGGA>C	10	0	0	10	100
76	11	120357842	85	Delezione di 2 bp	CTT>C	10	0	0	10	100
78	12	2834814	84	Inserzione di 21 bp	T>TTCTCAGTACGGTGAACCCAG	24	0	0	24	100
84	13	25817002	89	Inserzione di 19 bp	C>CAAAATATAAAAAGCTCCCT	24	0	0	24	100
85	13	44880152	89	Inserzione di 4 bp	C>CCTGT	12	0	0	12	100
86	13	77665265	77	Delezione di 20 bp	ATCTATTTTCTAATAGACGGC>A	14	0	0	14	100
89	14	46958967	73	Delezione di 22 bp	TTTAAATTTGAATGTGATAAAA>T	24	0	0	24	100
90	14	58050081	81	Inserzione di 4 bp	C>CTGAT	20	0	0	20	100
91	14	82390602	91	Delezione di 16 bp	CTTGCTCTATAAACCGT>C	10	0	0	10	100
93	14	102808554	94	Delezione di 5 bp	CGTGGA>C	10	0	0	10	100
95	15	63446199	68	Delezione di 6 bp	CAAATT>C	12	0	0	12	100
96	15	77879862	95	Delezione di 25 bp	GCCCCTGAGCCAGCCTCCCCTCTTA>G	14	0	0	14	100
98	15	85438311	72	Inserzione di 3 bp	C>CTTG	8	0	0	8	100
100	15	89864316	70	Inserzione di 4 bp	G>GCTAC	8	0	0	8	100

Amplicone	Cromosoma	Posizione	Dimensione del frammento analizzato	Tipo e lunghezza dell'indel per l'amplicone	Indel	N. totale di replicati del campione con indel	N. di identificazioni non riuscite	N. totale di identificazioni indel errate	N. totale di identificazioni indel corrette	PPA
105	16	85706416	91	Delezione di 7 bp	ATTATTTTC>A	16	0	0	16	100
107	17	3594276	87	Delezione di 1 bp	TG>T	2	0	0	2	100
108	17	3970133	91	Inserzione di 18 bp	A>ATCCTATTCTACTCTGAAT	10	0	0	10	100
109	17	16084985	93	Inserzione di 4 bp	A>AACAC	10	0	0	10	100
111	17	39589692	84	Inserzione di 1 bp	C>CA	16	9	7	0	0
112	17	39589739	84	Inserzione di 24 bp	T>TTCTGAAGGTCAAGTCTATCCCTGA	24	0	0	24	100
113	17	45438886	92	Delezione di 4 bp	CAGTG>C	12	0	0	12	100
114	17	61502459	79	Delezione di 12 bp	TTTGTATCTGCTG>T	20	0	0	20	100
120	18	38837054	75	Inserzione di 22 bp	T>TGTATCTTAGCAAAAGTTTCTCA	24	0	0	24	100
121	18	47405425	81	Inserzione di 3 bp	T>TGAG	20	0	0	20	100
122	18	54815706	85	Delezione di 2 bp	ACT>A	20	0	0	20	100
130	20	21766863	70	Delezione di 15 bp	TACTTGAGAACTGAGG>T	4	0	0	4	100
131	20	25278464	101	Inserzione di 5 bp	A>AGTGGG	20	0	0	20	100
132	20	50897361	67	Inserzione di 11 bp	G>GGAATGTCAGCC	24	0	0	24	100
134	20	62690925	87	Delezione di 16 bp	TCCTGGCTGGCCTGTGG>T	10	0	0	10	100
135	21	30300873	66	Inserzione di 11 bp	G>GATAAAACTTTA	10	0	0	10	100
137	21	36710749	87	Delezione di 21 bp	ACTCAAGATAACTCATGTTATC>A	16	0	0	16	100
138	21	46644985	69	Delezione di 5 bp	GTTGTT>G	8	0	0	8	100
140	22	25750814	100	Inserzione di 6 bp	C>CAGGGCA	20	0	0	20	100
142	22	37409885	97	Inserzione di 5 bp	C>CTGTTT	2	0	0	2	100
144	22	47081407	92	Delezione di 10 bp	GGGCACAGGCA>G	12	0	0	12	100

## Riproducibilità

Sono stati condotti due studi per valutare la riproducibilità dello strumento MiSeqDx con linee cellulari (studio 1 e 2) o sangue leucodepleto a cui sono state aggiunte linee cellulari (studio 2). Lo Studio 1 ha utilizzato più strumenti. Lo Studio 2 è stato svolto in più siti.

## Studio 1

La riproducibilità dello strumento MiSeqDx è stata determinata utilizzando due strumenti, due operatori e due lotti di reagenti, per un totale di otto corse. Il saggio rappresentativo, i campioni e il metodo di riferimento sono gli stessi di quelli descritti per lo studio di accuratezza.

I risultati sono presentati in base ai singoli ampliconi per ciascuno strumento [Tabella 6](#), per dimostrare la riproducibilità delle identificazioni tra strumenti. La percentuale di identificazioni corrette include le identificazioni errate e le identificazioni non riuscite (l'identificazione di varianti non soddisfa uno o più filtri). Gli strumenti hanno generato numeri simili di identificazioni non riuscite e identificazioni errate dovute al determinato amplicone.

Tabella 6 Risultati della riproducibilità tra strumenti dello strumento MiSeqDx (a livello di amplicone)

Amplicone	Cromosoma	Avvio amplicone	Fine amplicone	Dimensione del frammento analizzato	Basi nella regione affidabile	Contenuto genomico dell'amplicone	Contenuto in GC	MiSeqDx 1			MiSeqDx 2		
								Totale di identificazioni corrette	Totale di identificazioni errate	N. totale di identificazioni non riuscite	Totale di identificazioni corrette	Totale di identificazioni errate	N. totale di identificazioni non riuscite
1	1	36450499	36450591	93	93	Indel	0,22	8.928	0	0	8.928	0	0
2	1	109465122	109465200	79	79	PolyA (5), PolyC (5), indel	0,38	7.584	0	0	7.584	0	0
3	1	218353867	218353957	91	91	Indel	0,4	8.736	0	0	8.736	0	0
4	1	223906657	223906748	92	92	Indel	0,49	8.832	0	0	8.832	0	0
5	1	228526602	228526682	81	81	PolyG (5)	0,69	7.776	0	0	7.776	0	0
6	1	236372039	236372108	70	70	PolyT (10), indel	0,39	6.720	0	0	6.720	0	0
7	1	247812041	247812128	88	88	PolyA (5), CT (3), TAA(3), indel	0,27	8.448	0	0	8.448	0	0
8	2	55862774	55862863	90	90	Indel	0,28	8.640	0	0	8.640	0	0
9	2	87003930	87004009	80	80	Indel	0,38	7.680	0	0	7.680	0	0
10	2	177016721	177016805	85	81	N/D	0,65	7.775	1	0	7.775	1	0
11	2	186625727	186625801	75	75	PolyA (8)	0,35	7.200	0	0	7.200	0	0
12	2	190323504	190323591	88	88	PolyT (5)	0,42	8.448	0	0	8.448	0	0
13	2	200796740	200796826	87	87	PolyT (5), indel	0,31	8352	0	0	8352	0	0

Amplicone	Cromosoma	Avvio amplicone	Fine amplicone	Dimensione del frammento analizzato	Basi nella regione affidabile	Contenuto genomico dell'amplicone	Contenuto in GC	MiSeqDx 1			MiSeqDx 2		
								Totale di identificazioni corrette	Totale di identificazioni errate	N. totale di identificazioni non riuscite	Totale di identificazioni corrette	Totale di identificazioni errate	N. totale di identificazioni non riuscite
14	2	212245049	212245139	91	91	PolyT (5), PolyA (6), indel	0,3	8.736	0	0	8.736	0	0
15	2	228147052	228147144	93	93	N/D	0,43	8.928	0	0	8.928	0	0
16	2	235016350	235016422	73	73	PolyT (5), indel	0,42	7.008	0	0	7.008	0	0
17	3	4466229	4466321	93	93	AT(3), indel	0,27	8.761	0	167	8.760	0	168
18	3	46620561	46620643	83	83	N/D	0,43	7.968	0	0	7.968	0	0
19	3	49851331	49851400	70	70	CT(3), indel	0,49	6.720	0	0	6.720	0	0
20	3	189713161	189713248	88	88	PolyA (5), PolyT (5), PolyA (9), TG (3)	0,41	8.448	0	0	8.448	0	0
21	3	190106030	190106104	75	74	Indel	0,57	7.096	0	8	7.096	0	8
22	4	2233667	2233744	78	78	PolyA (6)	0,26	7.488	0	0	7.488	0	0
23	4	7780541	7780637	97	97	PolyG (6), PolyT (5), PolyA (5)	0,42	9.312	0	0	9.312	0	0
24	4	15688604	15688681	78	78	N/D	0,29	7.488	0	0	7.488	0	0
25	4	56236521	56236586	66	62	PolyA (5), indel	0,36	5.952	0	0	5.952	0	0
26	4	102839244	102839314	71	69	PolyA (5)	0,46	6.624	0	0	6.624	0	0
27	4	164446743	164446804	62	62	PolyA (7), indel	0,27	5.952	0	0	5.952	0	0
28	5	1882081	1882158	78	75	N/D	0,78	7.200	0	0	7.200	0	0
29	5	14769061	14769144	84	84	GT(3), CCA (3)	0,62	8.064	0	0	8.064	0	0
30	5	41069808	41069871	64	64	N/D	0,39	6.144	0	0	6.144	0	0
31	5	74077114	74077196	83	83	PolyA (6), indel	0,3	7.968	0	0	7.968	0	0
32	5	147475343	147475409	67	67	PolyT (5)	0,37	6.432	0	0	6.432	0	0

Amplicone	Cromosoma	Avvio amplicone	Fine amplicone	Dimensione del frammento analizzato	Basi nella regione affidabile	Contenuto genomico dell'amplicone	Contenuto in GC	MiSeqDx 1			MiSeqDx 2		
								Totale di identificazioni corrette	Totale di identificazioni errate	N. totale di identificazioni non riuscite	Totale di identificazioni corrette	Totale di identificazioni errate	N. totale di identificazioni non riuscite
33	5	149323731	149323821	91	91	CT(4), AG(3)	0,55	8.736	0	0	8.736	0	0
34	5	155662213	155662287	75	75	Indel	0,43	7.200	0	0	7.200	0	0
35	6	6318713	6318814	102	102	PolyG (6)	0,68	9.792	0	0	9.792	0	0
36	6	24949983	24950074	92	92	Indel	0,63	8.832	0	0	8.832	0	0
37	6	31084900	31084999	100	94	GCT(5), indel	0,61	8.979	0	45	8.979	0	45
38	6	32147987	32148084	98	98	PolyT (5), TCT(3), CTT (3)	0,55	9.408	0	0	9.408	0	0
39	6	32986864	32986958	95	95	Indel	0,53	9.120	0	0	9.120	0	0
40	6	33408498	33408583	86	86	PolyC (6)	0,7	8.256	0	0	8.256	0	0
41	6	41647401	41647495	95	94	PolyG (5), indel	0,61	9.024	0	0	9.024	0	0
42	6	112435865	112435955	91	91	PolyA (5)	0,44	8.736	0	0	8.736	0	0
43	7	22202076	22202148	73	73	N/D	0,44	7.008	0	0	7.008	0	0
44	7	66276100	66276187	88	88	Indel	0,35	8.448	0	0	8.448	0	0
45	7	77365735	77365821	87	87	PolyA (7), AG (4)	0,26	8352	0	0	8352	0	0
46	7	110939946	110940030	85	85	Indel	0,38	8.160	0	0	8.160	0	0
47	7	128533468	128533557	90	90	PolyG (5), indel	0,62	8.550	0	90	8.550	0	90
48	7	149503875	149503965	91	91	PolyG (6), PolyC (6), indel	0,71	8.736	0	0	8.736	0	0
49	7	154404519	154404599	81	66	N/D	0,31	6.336	0	0	6.336	0	0
50	7	156476507	156476599	93	93	Indel	0,35	8.928	0	0	8.928	0	0
51	8	1817312	1817394	83	83	N/D	0,42	7.968	0	0	7.968	0	0
52	8	24811020	24811109	90	89	PolyG (7), CTC(4), indel	0,61	8452	0	92	8.449	0	95
53	8	76518625	76518691	67	67	Indel	0,3	6.432	0	0	6.432	0	0
54	9	103054909	103055006	98	98	PolyG (6)	0,67	9.408	0	0	9.408	0	0

Amplicone	Cromosoma	Avvio amplicone	Fine amplicone	Dimensione del frammento analizzato	Basi nella regione affidabile	Contenuto genomico dell'amplicone	Contenuto in GC	MiSeqDx 1			MiSeqDx 2		
								Totale di identificazioni corrette	Totale di identificazioni errate	N. totale di identificazioni non riuscite	Totale di identificazioni corrette	Totale di identificazioni errate	N. totale di identificazioni non riuscite
55	9	105586150	105586214	65	65	Indel	0,32	6.240	0	0	6.240	0	0
56	9	107620823	107620918	96	96	N/D	0,49	9.216	0	0	9.216	0	0
57	9	123769149	123769231	83	83	AT(3)	0,37	7.968	0	0	7.968	0	0
58	9	138995345	138995441	97	97	PolyC (6), indel	0,68	9.312	0	0	9.312	0	0
59	10	5987120	5987198	79	78	PolyG (5), indel	0,47	7.488	0	0	7.488	0	0
60	10	11784629	11784726	98	91	GC(3)	0,87	8.644	1	91	8.644	1	91
61	10	27317777	27317855	79	79	PolyT (5)	0,3	7.584	0	0	7.584	0	0
62	10	33018351	33018440	90	90	PolyA (5), PolyT (5)	0,2	8.640	0	0	8.640	0	0
63	10	45084159	45084253	95	95	Indel	0,35	9.120	0	0	9.120	0	0
64	10	55892599	55892687	89	88	AC(11), indel	0,42	8.408	0	40	8.407	0	41
65	10	101611250	101611329	80	80	N/D	0,49	7.680	0	0	7.680	0	0
66	10	118351373	118351453	81	81	N/D	0,51	7.776	0	0	7.776	0	0
67	11	8159816	8159912	97	96	N/D	0,45	9.216	0	0	9.216	0	0
68	11	30177648	30177717	70	70	Indel	0,46	6.720	0	0	6.720	0	0
69	11	47470345	47470444	100	100	N/D	0,65	9.600	0	0	9.600	0	0
70	11	59837679	59837740	62	62	Indel	0,37	5.952	0	0	5.952	0	0
71	11	64418856	64418957	102	102	N/D	0,59	9.792	0	0	9.792	0	0
72	11	93529612	93529684	73	73	PolyA (5)	0,4	7.008	0	0	7.008	0	0
73	11	101347052	101347136	85	85	N/D	0,42	8.160	0	0	8.160	0	0
74	11	102477336	102477426	91	91	PolyG (6)	0,55	8.736	0	0	8.736	0	0
75	11	118406285	118406369	85	85	Indel	0,53	8.160	0	0	8.160	0	0
76	11	120357801	120357885	85	85	PolyA (5), CA (3), indel	0,34	8.160	0	0	8.160	0	0
77	11	125769313	125769397	85	85	GA(3)	0,52	8.160	0	0	8.160	0	0
78	12	2834770	2834853	84	84	PolyC (5), indel	0,52	8.064	0	0	8.064	0	0

Amplicone	Cromosoma	Avvio amplicone	Fine amplicone	Dimensione del frammento analizzato	Basi nella regione affidabile	Contenuto genomico dell'amplicone	Contenuto in GC	MiSeqDx 1			MiSeqDx 2		
								Totale di identificazioni corrette	Totale di identificazioni errate	N. totale di identificazioni non riuscite	Totale di identificazioni corrette	Totale di identificazioni errate	N. totale di identificazioni non riuscite
79	12	26811004	26811096	93	93	PolyA (7), AC (4)	0,33	8.928	0	0	8.928	0	0
80	12	30881766	30881846	81	81	N/D	0,49	7.776	0	0	7.776	0	0
81	12	88474105	88474175	71	71	PolyA (6)	0,35	6.816	0	0	6.816	0	0
82	12	120966872	120966966	95	95	PolyG (5)	0,68	9.117	3	0	9.119	1	0
83	13	24167504	24167576	73	73	N/D	0,52	7.008	0	0	7.008	0	0
84	13	25816961	25817049	89	88	PolyA (5), PolyT (7), PolyA (7), indel	0,22	8.448	0	0	8.448	0	0
85	13	44880112	44880200	89	89	Indel	0,49	8.544	0	0	8.544	0	0
86	13	77665218	77665294	77	77	Indel	0,39	7.392	0	0	7.392	0	0
87	14	31619327	31619393	67	67	GA(3), TA(3)	0,39	6.432	0	0	6.432	0	0
88	14	39517884	39517966	83	83	N/D	0,25	7.968	0	0	7.968	0	0
89	14	46958962	46959034	73	72	PolyT (5), indel	0,19	6.830	0	82	6.835	0	77
90	14	58050030	58050110	81	81	Indel	0,38	7.776	0	0	7.776	0	0
91	14	82390559	82390649	91	91	Indel	0,35	8.736	0	0	8.736	0	0
92	14	92549544	92549609	66	66	PolyA (5)	0,41	6.336	0	0	6.336	0	0
93	14	102808496	102808589	94	94	Indel	0,62	9.024	0	0	9.024	0	0
94	15	43170751	43170848	98	96	PolyC (5)	0,45	9.216	0	0	9.216	0	0
95	15	63446149	63446216	68	68	Indel	0,25	6.528	0	0	6.528	0	0
96	15	77879807	77879901	95	93	PolyG (5), indel	0,68	8.928	0	0	8.926	2	0
97	15	81625334	81625428	95	95	PolyT (6)	0,43	9.120	0	0	9.120	0	0
98	15	85438263	85438334	72	71	Indel	0,65	6.816	0	0	6.816	0	0
99	15	89817413	89817503	91	91	N/D	0,36	8.736	0	0	8.736	0	0
100	15	89864274	89864343	70	70	Indel	0,56	6.720	0	0	6.720	0	0
101	16	1894910	1894972	63	63	N/D	0,27	6.048	0	0	6.048	0	0

Amplicone	Cromosoma	Avvio amplicone	Fine amplicone	Dimensione del frammento analizzato	Basi nella regione affidabile	Contenuto genomico dell'amplicone	Contenuto in GC	MiSeqDx 1			MiSeqDx 2		
								Totale di identificazioni corrette	Totale di identificazioni errate	N. totale di identificazioni non riuscite	Totale di identificazioni corrette	Totale di identificazioni errate	N. totale di identificazioni non riuscite
102	16	28997904	28997998	95	95	PolyC (5)	0,67	9.120	0	0	9.120	0	0
103	16	53682908	53682994	87	87	TA(3)	0,41	8352	0	0	8352	0	0
104	16	57954406	57954509	104	104	PolyC (5)	0,67	9.984	0	0	9.984	0	0
105	16	85706375	85706465	91	91	PolyT (5), indel	0,37	8.736	0	0	8.736	0	0
106	17	3563920	3564008	89	89	GC(3)	0,64	8.544	0	0	8.544	0	0
107	17	3594191	3594277	87	87	PolyC (5), indel	0,67	8.347	0	5	8.347	0	5
108	17	3970090	3970180	91	91	Indel	0,46	8.736	0	0	8.736	0	0
109	17	16084945	16085037	93	93	Indel	0,26	8.928	0	0	8.928	0	0
110	17	33998759	33998849	91	89	PolyT (5)	0,54	8.544	0	0	8.544	0	0
111	17	39589691	39589774	84	82	PolyA (13), indel (x2)	0,29	7.776	7	89	7.777	12	83
112	17	41244394	41244484	91	91	PolyA (5)	0,34	8.736	0	0	8.736	0	0
113	17	45438866	45438957	92	92	PolyA (7), AT (3), AT(4), AT (4), indel	0,26	8.832	0	0	8.832	0	0
114	17	61502432	61502510	79	79	Indel	0,41	7.546	0	38	7.547	0	37
115	17	64023582	64023667	86	86	PolyT (7)	0,22	8.256	0	0	8.256	0	0
116	17	72308237	72308320	84	84	GAG(3)	0,62	8.064	0	0	8.064	0	0
117	18	2616456	2616522	67	67	GA(3)	0,31	6.432	0	0	6.432	0	0
118	18	6980478	6980568	91	91	N/D	0,37	8.736	0	0	8.736	0	0
119	18	9888026	9888094	69	69	PolyA (6), TG (3)	0,43	6.624	0	0	6.624	0	0
120	18	38836999	38837073	75	75	PolyA (5), indel	0,37	7.200	0	0	7.200	0	0
121	18	47405382	47405462	81	81	CTC(3), indel	0,47	7.776	0	0	7.776	0	0
122	18	54815665	54815749	85	85	CT(3), indel	0,45	8.160	0	0	8.160	0	0
123	18	59773996	59774060	65	65	N/D	0,48	6.240	0	0	6.240	0	0



Amplicone	Cromosoma	Avvio amplicone	Fine amplicone	Dimensione del frammento analizzato	Basi nella regione affidabile	Contenuto genomico dell'amplicone	Contenuto in GC	MiSeqDx 1			MiSeqDx 2		
								Totale di identificazioni corrette	Totale di identificazioni errate	N. totale di identificazioni non riuscite	Totale di identificazioni corrette	Totale di identificazioni errate	N. totale di identificazioni non riuscite
124	19	625143	625241	99	99	N/D	0,59	9.504	0	0	9.504	0	0
125	19	18121418	18121491	74	74	N/D	0,68	7.102	2	0	7.104	0	0
126	19	18186574	18186643	70	70	N/D	0,64	6.718	2	0	6.718	2	0
127	20	746056	746149	94	94	N/D	0,61	9.024	0	0	9.024	0	0
128	20	10633195	10633276	82	82	AC(3)	0,59	7.872	0	0	7.872	0	0
129	20	17705633	17705708	76	76	CT(3)	0,58	7.296	0	0	7.296	0	0
130	20	21766821	21766890	70	70	GT(3), TG(4), indel	0,46	6.720	0	0	6.720	0	0
131	20	25278421	25278521	101	101	Indel	0,63	9.696	0	0	9.696	0	0
132	20	50897302	50897368	67	67	Indel	0,36	6.432	0	0	6.432	0	0
133	20	62331904	62331994	91	88	PolyG (6)	0,73	8.360	0	88	8.360	0	88
134	20	62690860	62690946	87	87	Indel	0,57	8352	0	0	8352	0	0
135	21	30300823	30300888	66	66	Indel	0,35	6.336	0	0	6.336	0	0
136	21	33694176	33694273	98	98	PolyT (6), CA (3)	0,54	9.408	0	0	9.408	0	0
137	21	36710706	36710792	87	87	GT(3), indel	0,39	8352	0	0	8352	0	0
138	21	46644924	46644992	69	69	PolyA (6), AG (3), indel	0,32	6.603	0	21	6.601	0	23
139	21	46705575	46705664	90	90	PolyT (5), PolyA (6)	0,5	8.640	0	0	8.640	0	0
140	22	25750774	25750873	100	100	Indel	0,63	9.600	0	0	9.600	0	0
141	22	32439233	32439329	97	97	N/D	0,68	9.312	0	0	9.312	0	0
142	22	37409844	37409940	97	97	Indel	0,46	9.312	0	0	9.312	0	0
143	22	37637596	37637694	99	99	N/D	0,6	9.504	0	0	9.504	0	0
144	22	47081347	47081438	92	92	Indel	0,66	8.832	0	0	8.832	0	0
145	X	15870424	15870492	69	69	PolyT (5)	0,26	6.624	0	0	6.624	0	0
146	X	135288543	135288611	69	69	PolyC (5)	0,62	6.624	0	0	6.624	0	0
147	X	135290777	135290847	71	71	N/D	0,52	6.816	0	0	6.816	0	0
148	Y	2655397	2655461	65	0	N/D	0,55	0	0	0	0	0	0

Amplicone	Cromosoma	Avvio amplicone	Fine amplicone	Dimensione del frammento analizzato	Basi nella regione affidabile	Contenuto genomico dell'amplicone	Contenuto in GC	MiSeqDx 1			MiSeqDx 2		
								Totale di identificazioni corrette	Totale di identificazioni errate	N. totale di identificazioni non riuscite	Totale di identificazioni corrette	Totale di identificazioni errate	N. totale di identificazioni non riuscite
149	Y	2655519	2655609	91	0	N/D	0,48	0	0	0	0	0	0
150	Y	2655609	2655679	71	0	PolyA (5)	0,37	0	0	0	0	0	0

I risultati dello studio di riproducibilità sono stati analizzati in base all'operatore utilizzando la frequenza di varianti (Tabella 7). Questa analisi ha dimostrato che, per gli operatori, le frequenze delle varianti erano coerenti. Sono presentate le frequenze delle varianti medie alla deviazione standard di +/- 1.

Tabella 7 Risultati tra operatori per lo strumento MiSeqDx

Intervallo delle frequenze delle varianti	N. varianti univoche	N. totale di varianti analizzate dall'operatore 1	N. totale di varianti analizzate dall'operatore 2	Frequenza della variante media (DS) riportata dall'operatore 1	Frequenza della variante media (DS) riportata dall'operatore 2
Omozigote (0,70-1,00)	2.424	2.424	2.422	0,94 +/- 0,07	0,96 +/- 0,05
Eterozigote (0,20-0,70)	8.240	8.132	8.128	0,48 +/- 0,04	0,49 +/- 0,04

I risultati dello studio di riproducibilità per ogni campione sono riportati combinati per tutte le otto corse (Tabella 8). Il rilevamento è stato valutato per ciascun tipo di variante: SNV, inserzioni e delezioni, separatamente. Sono state escluse le posizioni dei riferimenti. Questa analisi dimostra che i risultati per le varianti erano riproducibili sui campioni.

Tabella 8 Concordezza dei risultati di identificazione delle basi dello strumento MiSeqDx per campione

Campione	SNV				Inserzioni				Delezioni			
	N. totale	N. totale TP	N. totale FP	N. totale FN	N. totale	N. totale TP	N. totale FP	N. totale FN	N. totale	N. totale TP	N. totale FP	N. totale FN
NA12877	592	592	0	0	336	336	0	0	288	288	0	0
NA12878	1.456	1.456	0	0	320	304	0	0	384	368	0	0
NA12879	912	912	0	0	336	320	0	2	288	288	0	0
NA12880	1.072	1.071	0	1	384	384	0	0	320	304	0	0

Campione	SNV				Inserzioni				Delezioni			
	N. totale	N. totale TP	N. totale FP	N. totale FN	N. totale	N. totale TP	N. totale FP	N. totale FN	N. totale	N. totale TP	N. totale FP	N. totale FN
NA12881	1.248	1.247	0	1	384	368	0	0	368	368	0	0
NA12882	944	943	0	1	352	336	0	4	304	288	0	0
NA12883	1.088	1.087	0	1	368	368	0	0	352	335	0	1
NA12884	1.088	1.088	0	0	400	384	0	5	336	336	0	0
NA12885	1.200	1.189	0	7	400	382	0	4	352	336	0	0
NA12886	1.104	1.102	0	2	368	352	0	3	368	368	0	0
NA12888	1.056	1.054	0	2	368	368	0	0	304	304	0	0
NA12893	1.168	1.168	0	0	352	336	0	1	368	368	0	0

I dati ottenuti nelle otto corse di questo studio di riproducibilità dimostrano che lo strumento MiSeqDx è in grado di sequenziare in modo coerente:

- contenuto in GC  $\geq 19\%$  (tutte le basi identificate in 192 dei 192 ampliconi sequenziati con 19% di contenuto in GC identificato correttamente con una percentuale di rilevamento non riuscito dell'1,1%)
- Contenuto in GC  $\leq 78\%$  (tutte le basi identificate in 192 dei 192 ampliconi sequenziati con il 78% di contenuto in GC identificato correttamente e nessun rilevamento non riuscito)
- Lunghezze PolyA  $\leq 8$  (la ripetizione PolyA di otto nucleotidi era stata identificata correttamente in 192 dei 192 ampliconi sequenziati contenenti PolyA = 8)
- Lunghezze PolyT  $\leq 10$  (la ripetizione PolyT di 10 nucleotidi era stata identificata correttamente in 192 dei 192 ampliconi sequenziati contenenti PolyT = 10)
- Lunghezze PolyG  $\leq 7$  (la ripetizione PolyG di 7 nucleotidi era stata identificata correttamente in 192 dei 192 ampliconi sequenziati contenenti PolyG =7)
- Lunghezze PolyC  $\leq 6$  (la ripetizione PolyC di 6 nucleotidi era stata identificata correttamente in 576 dei 576 ampliconi sequenziati contenenti PolyC =6)
- Lunghezze di dinucleotidi ripetuti  $\leq 11x$  (tutte le basi identificate in 192 dei 192 ampliconi sequenziati con 11x di ripetizione di dinucleotidi sono state identificate correttamente con una percentuale di rilevamento non riuscito dello 0,5%)
- Lunghezze di trinucleotidi ripetuti  $\leq 5x$  (tutte le basi identificate in 192 dei 192 ampliconi sequenziati con 5x di ripetizione di trinucleotide sono state identificate correttamente con una percentuale di rilevamento non riuscito dello 0,5%)

- 24 o meno inserzioni delle basi e 25 o meno delezioni della basi
  - 24 inserzioni delle basi identificate correttamente in 192 dei 192 campioni
  - 25 delezioni delle basi identificate correttamente in 223 campioni e un'identificazione non riuscita in uno su 224 campioni

## Studio 2

Uno studio di riproducibilità tra siti condotto utilizzando un saggio rappresentativo, il saggio Illumina MiSeqDx Cystic Fibrosis 139 Variant Assay, ha incluso un sottoinsieme di variazioni genetiche clinicamente significative del gene *CFTR*, analizzate con il software MiSeq Reporter usando il flusso di lavoro di sequenziamento del DNA bersaglio sulla piattaforma MiSeqDx. Lo studio in cieco usava tre siti per il trial e due operatori in ciascun laboratorio. Due pannelli ben caratterizzati di 46 campioni ciascuno sono stati testati da ciascun operatore in ciascun sito per un totale di 810 identificazioni per sito. I pannelli erano costituiti da una miscela di DNA genomico proveniente da linee cellulari con varianti note del gene *CFTR*, oltre che da sangue depleto di leucociti con aggiunta di linee cellulari con varianti note del gene *CFTR*. I campioni di sangue hanno consentito l'incorporazione delle fasi di estrazione necessarie per preparare il gDNA utilizzato come input primario per il flusso di lavoro del saggio. La percentuale dei campioni "pass", vale a dire il numero di campioni che hanno superato la metrica QC al primo tentativo, è stata del 99,88%. Tutti i risultati del test sono basati su test iniziali.

**Tabella 9** Riepilogo dei risultati dello studio di riproducibilità condotto utilizzando un saggio rappresentativo MiSeqDx Cystic Fibrosis 139-Variant Assay

Pannello	N. campione	Genotipo del campione	Varianti	Identificazioni totali per sede	Identificazioni con concordanza positiva (varianti)			Identificazioni con concordanza negativa (wild type)			N. di identificazioni mancate	N. di identificazioni non riuscite	Concordanza positiva (%)	Concordanza negativa (%)	Concordanza complessiva (%)
					Sede	Sede	Sede	Sede	Sede	Sede					
					1	2	3	1	2	3					
A	1	S549N (HET)		810	6	6	6	804	804	804	0	0	100	100	100
A	2	1812-1G>A (HET)		810	6	6	6	804	804	804	0	0	100	100	100
A	3	Q493X/F508del (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	4 <sup>1</sup>	F508del/2184delA (HET)		810	12	12	12	797	798	798	0	1 <sup>1</sup>	100	100	100
A	5 <sup>2</sup>	Y122X/R1158X (HET)		810	12	10	12	798	665	798	0	135 <sup>2</sup>	94,44	94,44	94,44
A	6	F508del/2183AA>G (HET)		810	12	12	12	798	798	798	0	0	100	100	100

Pannello	N. campione	Genotipo del campione	Varianti	Identificazioni totali per sede	Identificazioni con concordanza positiva (varianti)			Identificazioni con concordanza negativa (wild type)			N. di identificazioni mancate	N. di identificazioni non riuscite	Concordanza positiva (%)	Concordanza negativa (%)	Concordanza complessiva (%)
					Sede 1	Sede 2	Sede 3	Sede 1	Sede 2	Sede 3					
A	7	R75X (HET)		810	6	6	6	804	804	804	0	0	100	100	100
A	8	I507del/F508del (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	9 <sup>3</sup>	F508del/W1282X (HET)		810	12	11	12	798	797	798	2 <sup>3</sup>	0	97,22	99,96	99,92
A	10 <sup>3</sup>	F508del/3272-26A>G (HET)		810	12	11	12	798	797	798	2 <sup>3</sup>	0	97,22	99,96	99,92
A	11	F508del/3849+10kbC>T (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	12	621+1G>T/3120+1G>A (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	13	E60X/F508del (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	14	M1101K (HET)		810	6	6	6	804	804	804	0	0	100	100	100
A	15	M1101K (HOM)		810	6	6	6	804	804	804	0	0	100	100	100
A	16	F508del (HOM)	I506V, I507V, F508C non presente	828	6	6	6	822	822	822	0	0	100	100	100
A	17	F508del/3659delC (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	18	R117H/F508del (HET)	(TG)10(T)9/ (TG)12(T)5	816	18	18	18	798	798	798	0	0	100	100	100
A	19	621+1G>T/711+1G>T (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	20	G85E/621+1G>T (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	21	A455E/F508del (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	22	F508del/R560T (HET)		810	12	12	12	798	798	798	0	0	100	100	100

Pannello	N. campione	Genotipo del campione	Varianti	Identificazioni totali per sede	Identificazioni con concordanza positiva (varianti)			Identificazioni con concordanza negativa (wild type)			N. di identificazioni mancate	N. di identificazioni non riuscite	Concordanza positiva (%)	Concordanza negativa (%)	Concordanza complessiva (%)
					Sede 1	Sede 2	Sede 3	Sede 1	Sede 2	Sede 3					
A	23	F508del/Y1092X (C>A) (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	24	N1303K (HET)		810	6	6	6	804	804	804	0	0	100	100	100
A	25	G542X (HOM)		810	6	6	6	804	804	804	0	0	100	100	100
A	26	G542X (HET)		810	6	6	6	804	804	804	0	0	100	100	100
A	27	G551D/R553X (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	28	3849+10kbC>T (HOM)		810	6	6	6	804	804	804	0	0	100	100	100
A	29	WT		810	0	0	0	810	810	810	0	0	N/D	100	100
A	30	F508del (HET)		810	6	6	6	804	804	804	0	0	100	100	100
A	31	1717-1G>A (HET)		810	6	6	6	804	804	804	0	0	100	100	100
A	32	R1162X (HET)		810	6	6	6	804	804	804	0	0	100	100	100
A	33	R347P/G551D (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	34	R334W (HET)		810	6	6	6	804	804	804	0	0	100	100	100
A	35	WT		810	0	0	0	810	810	810	0	0	N/D	100	100
A	36	G85E (HET)		810	6	6	6	804	804	804	0	0	100	100	100
A	37	I336K (HET)		810	6	6	6	804	804	804	0	0	100	100	100
A	38	WT		810	0	0	0	810	810	810	0	0	N/D	100	100
A	39	F508del/3849+10kbC>T (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	40	621+1G>T/3120+1G>A (HET)		810	12	12	12	798	798	798	0	0	100	100	100

Pannello	N. campione	Genotipo del campione	Varianti	Identificazioni totali per sede	Identificazioni con concordanza positiva (varianti)			Identificazioni con concordanza negativa (wild type)			N. di identificazioni mancate	N. di identificazioni non riuscite	Concordanza positiva (%)	Concordanza negativa (%)	Concordanza complessiva (%)
					Sede 1	Sede 2	Sede 3	Sede 1	Sede 2	Sede 3					
A	41	F508del/3659delC (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	42	R117H/F508del (HET)	(TG)10(T)9/ (TG)12(T)5	816	18	18	18	798	798	798	0	0	100	100	100
A	43	G85E/621+1G>T (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	44	A455E/F508del (HET)		810	12	12	12	798	798	798	0	0	100	100	100
A	45	N1303K (HET)		810	6	6	6	804	804	804	0	0	100	100	100
A	46	G551D/R553X (HET)		810	12	12	12	798	798	798	0	0	100	100	100
B	47	2789+5G>A (HOM)		810	6	6	6	804	804	804	0	0	100	100	100
B	48	CFTR dele2, 3/F508del (HET)		810	12	12	12	798	798	798	0	0	100	100	100
B	49	F508del/1898+1G>A (HET)		810	12	12	12	798	798	798	0	0	100	100	100
B	50	WT		810	0	0	0	810	810	810	0	0	N/D	100	100
B	51	F508del/2143delT (HET)		810	12	12	12	798	798	798	0	0	100	100	100
B	52	3876delA (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	53	3905insT (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	54	394delTT (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	55	F508del (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	56	WT		810	0	0	0	810	810	810	0	0	N/D	100	100
B	57	WT		810	0	0	0	810	810	810	0	0	N/D	100	100
B	58	F508del (HET)		810	6	6	6	804	804	804	0	0	100	100	100

Pannello	N. campione	Genotipo del campione	Varianti	Identificazioni totali per sede	Identificazioni con concordanza positiva (varianti)			Identificazioni con concordanza negativa (wild type)			N. di identificazioni mancate	N. di identificazioni non riuscite	Concordanza positiva (%)	Concordanza negativa (%)	Concordanza complessiva (%)
					Sede 1	Sede 2	Sede 3	Sede 1	Sede 2	Sede 3					
B	59	WT		810	0	0	0	810	810	810	0	0	N/D	100	100
B	60	L206W (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	61	WT		810	0	0	0	810	810	810	0	0	N/D	100	100
B	62	G330X (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	63	WT		810	0	0	0	810	810	810	0	0	N/D	100	100
B	64	R347H (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	65	1078delT (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	66	G178R/F508del (HET)		810	12	12	12	798	798	798	0	0	100	100	100
B	67	S549R (c.1647T>G) (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	68	S549N (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	69	W846X (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	70	WT		810	0	0	0	810	810	810	0	0	N/D	100	100
B	71	E92X/F508del (HET)		810	12	12	12	798	798	798	0	0	100	100	100
B	72 <sup>4</sup>	621+1G>T/1154insTC (HET)		810	12	12	12	798	798	797	0	1 <sup>4</sup>	100	99,96	99,96
B	73	G542X (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	74	F508del (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	75 <sup>2</sup>	F508del (HET)		810	6	5	6	804	670	804	0	135 <sup>2</sup>	94,44	94,44	94,44
B	76	F508del (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	77	621+1G>T/A455E (HET)		810	12	12	12	798	798	798	0	0	100	100	100



Pannello	N. campione	Genotipo del campione	Varianti	Identificazioni totali per sede	Identificazioni con concordanza positiva (varianti)			Identificazioni con concordanza negativa (wild type)			N. di identificazioni mancate	N. di identificazioni non riuscite	Concordanza positiva (%)	Concordanza negativa (%)	Concordanza complessiva (%)
					Sede 1	Sede 2	Sede 3	Sede 1	Sede 2	Sede 3					
B	78	1812-1G>A (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	79	WT		810	0	0	0	810	810	810	0	0	N/D	100	100
B	80	F508del/R553X (HET)		810	12	12	12	798	798	798	0	0	100	100	100
B	81	F508del/G551D (HET)		810	12	12	12	798	798	798	0	0	100	100	100
B	82	R347P/F508del (HET)		810	12	12	12	798	798	798	0	0	100	100	100
B	83	R117H/F508del (HET)	(TG)10(T)9/ (TG)12(T)5	816	18	18	18	798	798	798	0	0	100	100	100
B	84	I507del (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	85	2789+5G>A (HOM)		810	6	6	6	804	804	804	0	0	100	100	100
B	86 <sup>4</sup>	CFTR dele2, 3/F508del (HET)		810	12	12	12	798	797	798	0	1 <sup>4</sup>	100	99,96	99,96
B	87	F508del/1898+1G>A (HET)		810	12	12	12	798	798	798	0	0	100	100	100
B	88	WT		810	0	0	0	810	810	810	0	0	N/D	100	100
B	89	F508del/2143delT (HET)		810	12	12	12	798	798	798	0	0	100	100	100
B	90	3905insT (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	91	394delTT (HET)		810	6	6	6	804	804	804	0	0	100	100	100
B	92	F508del (HET)		810	6	6	6	804	804	804	0	0	100	100	100
Totale				74.556	2.209			221.182			4	273	99,77	99,88	99,88

<sup>1</sup> La posizione del wild type corrispondente alla variante N1303K per un replicato ha prodotto un'identificazione non riuscita a causa della copertura insufficiente.

<sup>2</sup> Un replicato dei campioni 5 e 75 ha registrato una percentuale di identificazione dello 0%. Ulteriore investigazione indica che i campioni potrebbero non essere stati aggiunti alla piastra campioni prima della preparazione delle librerie perché i volumi dei campioni rimanenti nelle provette erano coerenti con nessun volume rimosso.

<sup>3</sup> Evidenze empiriche indicano che probabilmente i campioni 9 e 10 sono stati scambiati dall'operatore prima della preparazione delle librerie.

<sup>4</sup> La posizione del wild type corrispondente alla variante M1V per un replicato di ciascuno dei due campioni ha prodotto un'identificazione non riuscita a causa della copertura insufficiente.

## Caratteristiche delle prestazioni per le varianti somatiche

Gli studi qui descritti hanno utilizzato il modulo Somatic Variant per analizzare i dati del sequenziamento, fatta eccezione per quegli studi che hanno utilizzato un pannello di due geni dove è stato utilizzato un modulo specifico per il saggio.

### Accuratezza

Sono stati condotti tre studi per valutare l'accuratezza dello strumento MiSeqDx con DNA estratto da campioni FFPE.

#### Studio 1

Lo studio ha utilizzato un saggio rappresentativo progettato per interrogare diversi geni che coprono 12.588 basi su 23 cromosomi diversi utilizzando 150 ampliconi. Il cromosoma Y non contiene regioni affidabili e non è stato valutato. I cinque campioni univoci utilizzati in questo studio appartengono a una singola famiglia (due genitori e tre bambini). Questi campioni sono stati spesso sequenziati da più laboratori e metodologie di sequenziamento. Tre campioni sono femminili e due sono maschili. Per lo studio, tutti i campioni sono stati fissati in formalina e inclusi in paraffina prima dell'estrazione del DNA. Il campione GM12877 è stato diluito, al livello di DNA, con il campione GM12878 per creare GM12877-D e ottenere un set di varianti con frequenze prossime al 5% e al 10%. Ciascun campione è stato analizzato in duplicato, fatta eccezione per GM12877-D, che è stato analizzato con cinque replicati. L'accuratezza è stata determinata per SNV, inserzioni e delezioni confrontando i dati dello studio su un database di riferimenti ben caratterizzati. Le sequenze di riferimento del database (Platinum Genomes versione 2016-01) sono state derivate dalla combinazione di più metodologie di sequenziamento, dati disponibili pubblicamente e informazioni ereditarie. Le regioni genomiche affidabili sono state definite in base a questo metodo di riferimento, se non altrimenti specificato. In totale, i campioni sono stati analizzati otto volte. Le tabelle dimostrano che l'accuratezza si basava su dati ottenuti dalla prima corsa.

La [Tabella 10](#) contiene i dati dello studio con la concordanza positiva e negativa espressa in percentuale in base ai singoli campioni, dove i risultati delle varianti sono confrontati con un metodo di riferimento composito ben caratterizzato per i calcoli della PPA. Sono stati combinati i tre tipi di varianti (SNV, inserzioni e delezioni). Poiché il metodo di riferimento fornisce solo risultati per le varianti di singolo nucleotide e inserzioni/delezioni, i risultati delle basi non varianti sono confrontati con la sequenza di riferimento del genoma umano versione hg19, per i calcoli della NPA.

**Tabella 10** Concordanza dei risultati di identificazione delle basi dello strumento MiSeqDx con dati di riferimento di sei campioni ben caratterizzati

Campione	Percentuale di identificazione media	N. totale di varianti	N. totale di varianti TP	N. totale di varianti FN	N. totale di identificazioni TN	PPA	NPA	OPA
GM12877	98,7	152	147	0	23.719	100	100	100
GM12878	98,4	270	260	0	23.482	100	100	100
GM12879	98,7	192	186	0	23.744	100	100	100
GM12885	99,1	244	236	0	23.713	100	100	100
GM12886	98,7	230	226	0	23.652	100	100	100
GM12877-D <sup>1</sup>		675	650	0		100	100	100
GM12877-D <sup>2</sup>	98,4	155	155	0	57.608	100	100	100

<sup>1</sup> Varianti con una frequenza superiore al 20%.

<sup>2</sup> Varianti con una frequenza inferiore al 20%.

I 150 ampliconi sono stati progettati in modo da coprire contenuto genomico vario. Il contenuto in GC degli ampliconi rientrava nell'intervallo compreso tra il 26% e l'87%. Gli ampliconi presentavano inoltre un intervallo di ripetizioni di singolo nucleotide (ad es., PolyA, PolyT), dinucleotide e trinucleotide. Nel saggio sono stati utilizzati sei campioni univoci. I dati sono stati compilati in base ai singoli ampliconi ([Tabella 11](#)), allo scopo di determinare gli effetti del contenuto genomico sulla percentuale di identificazioni corrette. La percentuale di identificazioni corrette consiste nell'identificazione di varianti e riferimenti ed è inferiore al 100% in presenza di identificazioni errate o non riuscite. L'identificazione non riuscita si verifica quando l'identificazione di varianti non soddisfa uno o più filtri (ad es., copertura insufficiente). Non sono state rilevate identificazioni errate. Il numero di identificazioni non riuscite variava considerevolmente tra gli ampliconi. Il contenuto in GC e diverse interazioni con il contenuto in GC erano i predittori più significativi delle identificazioni non riuscite. 2.040/2.580 (79%) delle identificazioni non riuscite non hanno soddisfatto la specifica di copertura. Gli ampliconi con contenuto in GC superiore al 78% hanno costituito la maggior parte delle identificazioni non riuscite. Un amplicone rappresentativo con il 78% di contenuto in GC ha presentato un totale di 675 identificazioni non riuscite. Un amplicone rappresentativo con l'87% ha presentato un totale di 1.365 identificazioni non riuscite. La copertura può essere incrementata riducendo il numero di campioni caricati sulla cella a flusso, che consente il rilevamento sugli ampliconi con contenuto in GC elevato.

Tabella 11 Dati di accuratezza a livello di amplicone

Amplicone	Cromosoma	Avvio amplicone	Fine amplicone	Dimensione del frammento analizzato	Basi nella regione affidabile	Contenuto genomico dell'amplicone	Contenuto in GC	N. totale di identificazioni corrette	N. totale di identificazioni errate	N. totale di identificazioni non riuscite	% di identificazioni corrette
1	1	36450499	36450591	93	93	Indel	0,22	1.395	0	0	100
2	1	109465122	109465200	79	79	PolyA (5), PolyC (5), indel	0,38	1.185	0	0	100
3	1	218353867	218353957	91	91	Indel	0,4	1.364	0	1	99,9
4	1	223906657	223906748	92	92	Indel	0,49	1.380	0	0	100
5	1	228526602	228526682	81	81	PolyG (5)	0,69	1.215	0	0	100
6	1	236372039	236372108	70	70	PolyT (10), indel	0,39	1050	0	0	100
7	1	247812041	247812128	88	88	PolyA (5), CT(3), TAA(3), indel	0,27	1.320	0	0	100
8	2	55862774	55862863	90	90	Indel	0,28	1.350	0	0	100
9	2	87003930	87004009	80	80	Indel	0,38	1.200	0	0	100
10	2	177016721	177016805	85	81	N/D	0,65	1.215	0	0	100
11	2	186625727	186625801	75	75	PolyA (8)	0,35	1.117	0	10	99,1
12	2	190323504	190323591	88	88	PolyT (5)	0,42	1.320	0	0	100
13	2	200796740	200796826	87	87	PolyT (5), indel	0,31	1.302	0	8	99,4
14	2	212245049	212245139	91	91	PolyT (5), PolyA (6), indel	0,3	1.365	0	0	100
15	2	228147052	228147144	93	93	N/D	0,43	1.395	0	0	100
16	2	235016350	235016422	73	73	PolyT (5), indel	0,42	1.095	0	0	100
17	3	4466229	4466321	93	93	AT(3), indel	0,27	1.349	0	46	96,7
18	3	46620561	46620643	83	83	N/D	0,43	1.245	0	0	100
19	3	49851331	49851400	70	70	CT(3), indel	0,49	1050	0	0	100
20	3	189713161	189713248	88	88	PolyA (5), PolyT (5), PolyA (9), TG(3)	0,41	1.305	0	30	97,8
21	3	190106030	190106104	75	74	Indel	0,57	1108	0	2	99,8
22	4	2233667	2233744	78	78	PolyA (6)	0,26	1.170	0	0	100
23	4	7780541	7780637	97	97	PolyG (6), PolyT (5), PolyA (5)	0,42	1.455	0	0	100
24	4	15688604	15688681	78	78	N/D	0,29	1.169	0	1	99,9
25	4	56236521	56236586	66	62	PolyA (5), indel	0,36	930	0	0	100
26	4	102839244	102839314	71	69	PolyA (5)	0,46	1.035	0	0	100

Amplicone	Cromosoma	Avvio amplicone	Fine amplicone	Dimensione del frammento analizzato	Basi nella regione affidabile	Contenuto genomico dell'amplicone	Contenuto in GC	N. totale di identificazioni corrette	N. totale di identificazioni errate	N. totale di identificazioni non riuscite	% di identificazioni corrette
27	4	164446743	164446804	62	62	PolyA (7), indel	0,27	920	0	10	98,9
28	5	1882081	1882158	78	75	N/D	0,78	450	0	675	40,0
29	5	14769061	14769144	84	84	GT(3), CCA(3)	0,62	1.260	0	0	100
30	5	41069808	41069871	64	64	N/D	0,39	960	0	0	100
31	5	74077114	74077196	83	83	PolyA (6), indel	0,3	1.245	0	0	100
32	5	147475343	147475409	67	67	PolyT (5)	0,37	1.005	0	0	100
33	5	149323731	149323821	91	91	CT(4), AG(3)	0,55	1.365	0	0	100
34	5	155662213	155662287	75	75	Indel	0,43	1.125	0	0	100
35	6	6318713	6318814	102	102	PolyG (6)	0,68	1.530	0	0	100
36	6	24949983	24950074	92	92	Indel	0,63	1.380	0	0	100
37	6	31084900	31084999	100	94	GCT(5), indel	0,61	1.383	0	27	98,1
38	6	32147987	32148084	98	98	PolyT (5), TCT(3), CTT(3)	0,55	1.455	0	15	99,0
39	6	32986864	32986958	95	95	Indel	0,53	1.425	0	0	100
40	6	33408498	33408583	86	86	PolyC (6)	0,7	1.290	0	0	100
41	6	41647401	41647495	95	94	PolyG (5), indel	0,61	1.410	0	0	100
42	6	112435865	112435955	91	91	PolyA (5)	0,44	1.365	0	0	100
43	7	22202076	22202148	73	73	N/D	0,44	1.095	0	0	100
44	7	66276100	66276187	88	88	Indel	0,35	1.320	0	0	100
45	7	77365735	77365821	87	87	PolyA (7), AG(4)	0,26	1.299	0	6	99,5
46	7	110939946	110940030	85	85	Indel	0,38	1.275	0	0	100
47	7	128533468	128533557	90	90	PolyG (5), indel	0,62	1.350	0	0	100
48	7	149503875	149503965	91	91	PolyG (6), PolyC (6), indel	0,71	1.365	0	0	100
49	7	154404519	154404599	81	66	N/D	0,31	990	0	0	100
50	7	156476507	156476599	93	93	Indel	0,35	1.395	0	0	100
51	8	1817312	1817394	83	83	N/D	0,42	1.245	0	0	100
52	8	24811020	24811109	90	89	PolyG (7), CTC(4), indel	0,61	1.305	0	30	97,8
53	8	76518625	76518691	67	67	Indel	0,3	1.005	0	0	100
54	9	103054909	103055006	98	98	PolyG (6)	0,67	1.470	0	0	100

Amplicone	Cromosoma	Avvio amplicone	Fine amplicone	Dimensione del frammento analizzato	Basi nella regione affidabile	Contenuto genomico dell'amplicone	Contenuto in GC	N. totale di identificazioni corrette	N. totale di identificazioni errate	N. totale di identificazioni non riuscite	% di identificazioni corrette
55	9	105586150	105586214	65	65	Indel	0,32	973	0	2	99,8
56	9	107620823	107620918	96	96	N/D	0,49	1.440	0	0	100
57	9	123769149	123769231	83	83	AT(3)	0,37	1.242	0	3	99,8
58	9	138995345	138995441	97	97	PolyC (6), indel	0,68	1.455	0	0	100
59	10	5987120	5987198	79	78	PolyG (5), indel	0,47	1.170	0	0	100
60	10	11784629	11784726	98	91	GC(3)	0,87	0	0	1.365	0
61	10	27317777	27317855	79	79	PolyT (5)	0,3	1.185	0	0	100
62	10	33018351	33018440	90	90	PolyA (5), PolyT (5)	0,2	1.350	0	0	100
63	10	45084159	45084253	95	95	Indel	0,35	1.425	0	0	100
64	10	55892599	55892687	89	88	AC(11), indel	0,42	1.290	0	69	94,9
65	10	101611250	101611329	80	80	N/D	0,49	1.200	0	0	100
66	10	118351373	118351453	81	81	N/D	0,51	1.215	0	0	100
67	11	8159816	8159912	97	96	N/D	0,45	1.440	0	0	100
68	11	30177648	30177717	70	70	Indel	0,46	1050	0	0	100
69	11	47470345	47470444	100	100	N/D	0,65	1.500	0	0	100
70	11	59837679	59837740	62	62	Indel	0,37	930	0	0	100
71	11	64418856	64418957	102	102	N/D	0,59	1.530	0	0	100
72	11	93529612	93529684	73	73	PolyA (5)	0,4	1.095	0	0	100
73	11	101347052	101347136	85	85	N/D	0,42	1.275	0	0	100
74	11	102477336	102477426	91	91	PolyG (6)	0,55	1.365	0	0	100
75	11	118406285	118406369	85	85	Indel	0,53	1.275	0	0	100
76	11	120357801	120357885	85	85	PolyA (5), CA(3), indel	0,34	1.275	0	0	100
77	11	125769313	125769397	85	85	GA(3)	0,52	1.275	0	0	100
78	12	2834770	2834853	84	84	PolyC (5), indel	0,52	1.260	0	14	98,9
79	12	26811004	26811096	93	93	PolyA (7), AC(4)	0,33	1.395	0	0	100
80	12	30881766	30881846	81	81	N/D	0,49	1.215	0	0	100
81	12	88474105	88474175	71	71	PolyA (6)	0,35	1.065	0	0	100
82	12	120966872	120966966	95	95	PolyG (5)	0,68	1.425	0	0	100

Amplicone	Cromosoma	Avvio amplicone	Fine amplicone	Dimensione del frammento analizzato	Basi nella regione affidabile	Contenuto genomico dell'amplicone	Contenuto in GC	N. totale di identificazioni corrette	N. totale di identificazioni errate	N. totale di identificazioni non riuscite	% di identificazioni corrette
83	12	24167504	24167576	73	73	N/D	0,52	1.095	0	0	100
84	13	25816961	25817049	89	88	PolyA (5), PolyT (7), PolyA (7), indel	0,22	1.305	0	15	98,9
85	13	44880112	44880200	89	89	Indel	0,49	1.335	0	0	100
86	13	77665218	77665294	77	77	Indel	0,39	1.155	0	0	100
87	14	31619327	31619393	67	67	GA(3), TA(3)	0,39	1.005	0	0	100
88	14	39517884	39517966	83	83	N/D	0,25	1.245	0	0	100
89	14	46958962	46959034	73	72	PolyT (5), indel	0,19	1.038	0	42	96,1
90	14	58050030	58050110	81	81	Indel	0,38	1.215	0	0	100
91	14	82390559	82390649	91	91	Indel	0,35	1.365	0	0	100
92	14	92549544	92549609	66	66	PolyA (5)	0,41	975	0	60	94,2
93	14	102808496	102808589	94	94	Indel	0,62	1.410	0	0	100
94	15	43170751	43170848	98	96	PolyC (5)	0,45	1.440	0	0	100
95	15	63446149	63446216	68	68	Indel	0,25	1.020	0	0	100
96	15	77879807	77879901	95	93	PolyG (5), indel	0,68	1.395	0	0	100
97	15	81625334	81625428	95	95	PolyT (6)	0,43	1.425	0	0	100
98	15	85438263	85438334	72	71	Indel	0,65	1.065	0	0	100
99	15	89817413	89817503	91	91	N/D	0,36	1.365	0	0	100
100	15	89864274	89864343	70	70	Indel	0,56	1050	0	0	100
101	16	1894910	1894972	63	63	N/D	0,27	945	0	0	100
102	16	28997904	28997998	95	95	PolyC (5)	0,67	1.425	0	0	100
103	16	3682908	53682994	87	87	TA(3)	0,41	1.305	0	0	100
104	16	57954406	57954509	104	104	PolyC (5)	0,67	1.560	0	0	100
105	16	85706375	85706465	91	91	Poly T (5), indel	0,37	1.362	0	3	99,8
106	17	3563920	3564008	89	89	GC(3)	0,64	1.335	0	0	100
107	17	3594191	3594277	87	87	PolyC (5), indel	0,67	1.303	0	2	99,8
108	17	3970090	3970180	91	91	Indel	0,46	1.365	0	0	100
109	17	16084945	16085037	93	93	Indel	0,26	1.395	0	0	100

Amplicone	Cromosoma	Avvio amplicone	Fine amplicone	Dimensione del frammento analizzato	Basi nella regione affidabile	Contenuto genomico dell'amplicone	Contenuto in GC	N. totale di identificazioni corrette	N. totale di identificazioni errate	N. totale di identificazioni non riuscite	% di identificazioni corrette
110	17	33998759	33998849	91	89	PolyT (5)	0,54	1.335	0	0	100
111	17	39589691	39589774	84	82	PolyA (13), indel (x2)	0,29	1.215	0	78	94,0
112	17	41244394	41244484	91	91	PolyA (5)	0,34	1.365	0	0	100
113	17	45438866	45438957	92	92	PolyA (7), AT(3), AT(4), AT(4), indel	0,26	1.365	0	15	98,9
114	17	61502432	61502510	79	79	Indel	0,41	1.175	0	10	99,2
115	17	64023582	64023667	86	86	PolyT (7)	0,22	1.289	0	1	99,9
116	17	72308237	72308320	84	84	GAG(3)	0,62	1.260	0	0	100
117	18	2616456	2616522	67	67	GA(3)	0,31	1.005	0	0	100
118	18	6980478	6980568	91	91	N/D	0,37	1.365	0	0	100
119	18	9888026	9888094	69	69	PolyA (6), TG(3)	0,43	1.035	0	0	100
120	18	38836999	38837073	75	75	PolyA (5), indel	0,37	1.121	0	19	98,3
121	18	47405382	47405462	81	81	CTC(3), indel	0,47	1.215	0	0	100
122	18	54815665	54815749	85	85	CT(3), indel	00,45	1.275	0	0	100
123	18	59773996	59774060	65	65	N/D	0,48	975	0	0	100
124	19	625143	625241	99	99	N/D	0,59	1.478	0	7	99,5
125	19	18121418	18121491	74	74	N/D	0,68	1.110	0	0	100
126	19	18186574	18186643	70	70	N/D	0,64	1050	0	0	100
127	20	746056	746149	94	94	N/D	0,61	1.410	0	0	100
128	20	10633195	10633276	82	82	AC(3)	0,59	1.230	0	0	100
129	20	17705633	17705708	76	76	CT(3)	0,58	1.140	0	0	100
130	20	21766821	21766890	70	70	GT(3), TG(4), indel	0,46	1050	0	0	100
131	20	25278421	25278521	101	101	Indel	0,63	1.515	0	0	100
132	20	50897302	50897368	67	67	Indel	0,36	1.005	0	6	99,4
133	20	62331904	62331994	91	88	Poly G (6)	0,73	1.320	0	0	100
134	20	62690860	62690946	87	87	Indel	0,57	1.305	0	0	100
135	21	30300823	30300888	66	66	Indel	0,35	990	0	0	100
136	21	33694176	33694273	98	98	PolyT (6), CA(3)	0,54	1.470	0	0	100



Amplicone	Cromosoma	Avvio amplicone	Fine amplicone	Dimensione del frammento analizzato	Basi nella regione affidabile	Contenuto genomico dell'amplicone	Contenuto in GC	N. totale di identificazioni corrette	N. totale di identificazioni errate	N. totale di identificazioni non riuscite	% di identificazioni corrette
137	21	36710706	36710792	87	87	GT(3), indel	0,39	1.305	0	0	100
138	21	46644924	46644992	69	69	PolyA (6), AG(3), indel	0,32	1029	0	7	99,3
139	21	46705575	46705664	90	90	PolyT (5), PolyA (6)	0,5	1.350	0	0	100
140	22	25750774	25750873	100	100	Indel	0,63	1.500	0	1	99,9
141	22	32439233	32439329	97	97	N/D	0,68	1.455	0	0	100
142	22	37409844	37409940	97	97	Indel	0,46	1.455	0	0	100
143	22	37637596	37637694	99	99	N/D	0,6	1.485	0	0	100
144	22	47081347	47081438	92	92	Indel	0,66	1.380	0	0	100
145	X	15870424	15870492	69	69	PolyT (5)	0,26	1.035	0	0	100
146	X	135288543	135288611	69	69	PolyC (5)	0,62	1.035	0	0	100
147	X	135290777	135290847	71	71	N/D	0,52	1.065	0	0	100
148	Y	2655397	2655461	65	0	N/D	0,55	0	0	0	N/D
149	Y	2655519	2655609	91	0	N/D	0,48	0	0	0	N/D
150	Y	2655609	2655679	71	0	PolyA (5)	0,37	0	0	0	N/D

Le varianti con identificazione non riuscita sono riepilogate nella [Tabella 12](#). I filtri che hanno fornito le identificazioni non riuscite sono elencati nella tabella.

**Tabella 12** Riepilogo delle varianti con identificazione non riuscita

N. amplicone	Crom:Pos	Variante	Contenuto dell'amplicone corrispondente	Filtro	Varianti non identificate	Varianti previste
28	5:1882129	T > G	78% GC	LowDP <sup>1</sup>	8	13
52	8:24811064	AG > A	PolyG (7), CTC(4), 61% GC	R3x6 <sup>2</sup>	15	15
60	10:11784633	C > T	PolyGC (3), 87% GC	LowDP	13	13
64	10:55892600	TAC > T	AC(11), 42% GC	R3x6	9	9
111	17:39589692	C > CA	PolyA (13), 29% GC	R3x6	13	13

<sup>1</sup> LowDP: bassa copertura. Una variante viene filtrata se la profondità in almeno uno dei pool in quella determinata posizione è inferiore a 900.

<sup>2</sup> R3x6: ripetere il filtro. Una variante viene filtrata se tutta o parte della variante è presente ripetutamente nel genoma di riferimento adiacente alla posizione della variante. Sono richieste almeno sei ripetizioni nel riferimento e sono prese in considerazione solo le ripetizioni con lunghezza fino a 3 bp.

I risultati del sequenziamento per il campione sono stati confrontati con un genotipo altamente affidabile per NA12878 dal National Institutes of Standards and Technology (NIST) (v.2.19). Dei 150 ampliconi, 92 ampliconi sono rientrati completamente nelle regioni genomiche altamente affidabili, 41 ampliconi presentavano una sovrapposizione parziale e 17 ampliconi non presentavano alcuna sovrapposizione nella sequenza NIST. Questo ha fornito 10.000 coordinate per replicato per il confronto. Le identificazioni delle basi non varianti sono state confrontate con la sequenza di riferimento del genoma umano versione hg19. I risultati dell'accuratezza sono mostrati nella [Tabella 13](#).

**Tabella 13** Concordanza dei risultati di identificazione delle basi dello strumento MiSeqDx con il riferimento NIST per il campione GM12878

Campione	N. di ampliconi	Percentuale di identificazione media	N. totale di identificazioni delle varianti TP	N. totale di identificazioni delle varianti FN	N. totale di identificazioni TN	N. totale di identificazioni FP	PPA	NPA	OPA
GM12878	150	98,43	206	0	19.231	0	100	100	100

I cinque campioni non diluiti sono stati ulteriormente analizzati per identificare piccole inserzioni e delezioni (indel) ([Tabella 14](#)). In alcuni casi, l'indel era comune tra due o più campioni, come illustrato nella colonna "N. totale di replicati del campione con indel". I risultati per entrambi i replicati dei cinque campioni sono inclusi nella [Tabella 14](#). Erano presenti in totale 71 indel di dimensioni variabili tra 1 e 24 bp per le inserzioni e 1 e 25 bp per le delezioni. Sono state identificate 68 indel, ciascuna con una percentuale di concordanza positiva di 1. Per tre inserzioni e delezioni l'identificazione corretta non è stata possibile perché ciascuna di queste varianti è risultata un'identificazione non riuscita a causa del filtro R3x6. Di conseguenza, la percentuale di concordanza positiva (PPA), che esclude le identificazioni non riuscite, non è stata calcolata. Le tre varianti erano una delezione di 1 bp (chr8 24811064 AG>A), una delezione di 2 bp (chr10 55892600 TAC>T) e un'inserzione di 1 bp (chr17 39589692 C>CA).

**Tabella 14** Riepilogo del rilevamento delle indel con lo strumento MiSeqDx

Amplicone	Cromosoma	Posizione	Dimensione del frammento analizzato	Tipo e lunghezza dell'indel per l'amplicone	Indel	N. totale di replicati del campione con indel	N. di identificazioni non riuscite	N. totale di identificazioni indel errate	N. totale di identificazioni indel corrette	PPA
1	1	36450544	93	Delezione di 25 bp	GAAAATTTAATGAAACACATTGCCT>G	7	0	0	7	100
2	1	109465165	79	Delezione di 3 bp	ACTT>A	9	0	0	9	100
3	1	218353908	91	Inserzione di 23 bp	T>TTTTAATAGCAAAAAGAGGCTAGA	15	0	0	15	100
4	1	223906701	92	Delezione di 17 bp	GACAGACTGTGAGGAAGA>G	11	0	0	11	100
6	1	236372081	70	Inserzione di 5 bp	C>CTTAAG	9	0	0	9	100
7	1	247812083	88	Inserzione di 3 bp	C>CATG	9	0	0	9	100
8	2	55862804	90	Inserzione di 7 bp	T>TTTGTTAA	13	0	0	13	100

Amplicone	Cromosoma	Posizione	Dimensione del frammento analizzato	Tipo e lunghezza dell'indel per l'amplicone	Indel	N. totale di replicati del campione con indel	N. di identificazioni non riuscite	N. totale di identificazioni indel errate	N. totale di identificazioni indel corrette	PPA
9	2	87003972	80	Delezione di 6 bp	TTATCTC>T	11	0	0	11	100
13	2	200796749	87	Inserzione di 5 bp	T>TTAAAA	15	0	0	15	100
14	2	212245090	91	Inserzione di 12 bp	C>CTGAAAATAGGAT	11	0	0	11	100
16	2	235016388	73	Inserzione di 2 bp	A>ATG	9	0	0	9	100
17	3	4466274	93	Delezione di 23 bp	TAACCTAAAATTACAAAATAACCC>T	13	0	0	13	100
19	3	49851375	70	Inserzione di 9 bp	C>CCTGGCTCCT	7	0	0	7	100
21	3	190106071	75	Delezione di 1 bp	AG>A	13	0	0	13	100
25	4	56236567	66	Delezione di 8 bp	TAACCGAAA>T	9	0	0	9	100
27	4	164446785	62	Inserzione di 11 bp	T>TTATGGTATTGA	9	0	0	9	100
31	5	74077155	83	Delezione di 4 bp	TAGTA>T	7	0	0	7	100
34	5	155662255	75	Inserzione di 8 bp	G>GCCTACTGA	13	0	0	13	100
36	6	24950035	92	Delezione di 21 bp	CCCTGGGTGCTATAGCCCACCA>C	11	0	0	11	100
37	6	31084942	100	Delezione di 3 bp	GCTT>G	15	0	0	15	100
39	6	32986905	95	Delezione di 25 bp	CTTTCACCTTCCCGTCTCATGCCAAG>C	7	0	0	7	100
41	6	41647442	95	Delezione di 23 bp	GGCATGAGGCTTGGTGACATGGCA>G	11	0	0	11	100
44	7	66276142	88	Inserzione di 1 bp	C>CT	13	0	0	13	100
46	7	110939983	85	Delezione di 4 bp	CAAGT>C	13	0	0	13	100
47	7	128533514	90	Inserzione di 1 bp	T>TC	15	0	0	15	100
48	7	149503916	91	Delezione di 4 bp	GGATA>G	7	0	0	7	100
50	7	156476548	93	Delezione di 11 bp	GAATCTGCACTT>G	13	0	0	13	100
52	8	24811064	90	Delezione di 1 bp	AG>A	15	15	0	0	N/D
53	8	76518677	67	Inserzione di 4 bp	T>TACTG	9	0	0	9	100
55	9	105586193	65	Inserzione di 4 bp	C>CAATT	13	0	0	13	100
58	9	138995370	97	Delezione di 21 bp	TCTGGGGGGCAGCCCCCTGAGGG>T	9	0	0	9	100
59	10	5987158	79	Delezione di 3 bp	TAAC>T	11	0	0	11	100
63	10	45084202	95	Delezione di 16 bp	AGCGTCTATAACCAAAT>A	11	0	0	11	100
64	10	55892600	89	Delezione di 2 bp	TAC>T	9	9	0	0	100
68	11	30177690	70	Inserzione di 2 bp	C>CTG	7	0	0	7	100

Amplicone	Cromosoma	Posizione	Dimensione del frammento analizzato	Tipo e lunghezza dell'indel per l'amplicone	Indel	N. totale di replicati del campione con indel	N. di identificazioni non riuscite	N. totale di identificazioni indel errate	N. totale di identificazioni indel corrette	PPA
70	11	59837721	62	Inserzione di 8 bp	T>TTATGAAAA	11	0	0	11	100
75	11	118406328	85	Delezione di 8 bp	CAGTGTGGA>C	9	0	0	9	100
76	11	120357842	85	Delezione di 2 bp	CTT>C	11	0	0	11	100
78	12	2834814	84	Inserzione di 21 bp	T>TTCTCAGTACGGTGAACCCAG	15	0	0	15	100
84	13	25817002	89	Inserzione di 19 bp	C>CAAAATATAAAAAGCTCCCT	15	0	0	15	100
85	13	44880152	89	Inserzione di 4 bp	C>CCTGT	11	0	0	11	100
86	13	77665265	77	Delezione di 20 bp	ATCTATTTTCTAATAGACGGC>A	9	0	0	9	100
89	14	46958967	73	Delezione di 22 bp	TTTAAAATTTGAATGTGATAAAA>T	15	0	0	15	100
90	14	58050081	81	Inserzione di 4 bp	C>CTGAT	13	0	0	13	100
91	14	82390602	91	Delezione di 16 bp	CTTGCTCTATAAACCGT>C	11	0	0	11	100
93	14	102808554	94	Delezione di 5 bp	CGTGGA>C	9	0	0	9	100
95	15	63446199	68	Delezione di 6 bp	CAAATT>C	11	0	0	11	100
96	15	77879862	95	Delezione di 25 bp	GCCCCTGAGCCAGCCTCCCCTCTTA>G	9	0	0	9	100
98	15	85438311	72	Inserzione di 3 bp	C>CTTG	9	0	0	9	100
100	15	89864316	70	Inserzione di 4 bp	G>GCTAC	9	0	0	9	100
105	16	85706416	91	Delezione di 7 bp	ATTATTC>A	11	0	0	11	100
107	17	3594276	87	Delezione di 1 bp	TG>T	13	0	0	13	100
108	17	3970133	91	Inserzione di 18 bp	A>ATCCTATTCTACTCTGAAT	11	0	0	11	100
109	17	16084985	93	Inserzione di 4 bp	A>AACAC	7	0	0	7	100
111	17	39589692	84	Inserzione di 1 bp	C>CA	13	13	0	0	100
112	17	39589739	84	Inserzione di 24 bp	T>TTCTGAAGGTCAAGTCTATCCCTGA	15	0	0	15	100
113	17	45438886	92	Delezione di 4 bp	CAGTG>C	7	0	0	7	100
114	17	61502459	79	Delezione di 12 bp	TTTGATCTGCTG>T	13	0	0	13	100
120	18	38837054	75	Inserzione di 22 bp	T>TGATCTTAGCAAAAAGTTTCTCA	15	0	0	15	100
121	18	47405425	81	Inserzione di 3 bp	T>TGAG	11	0	0	11	100
122	18	54815706	85	Delezione di 2 bp	ACT>A	13	0	0	13	100
130	20	21766863	70	Delezione di 15 bp	TACTTGAGAACTGAGG>T	9	0	0	9	100
131	20	25278464	101	Inserzione di 5 bp	A>AGTGGG	13	0	0	13	100

Amplicone	Cromosoma	Posizione	Dimensione del frammento analizzato	Tipo e lunghezza dell'indel per l'amplicone	Indel	N. totale di replicati del campione con indel	N. di identificazioni non riuscite	N. totale di identificazioni indel errate	N. totale di identificazioni indel corrette	PPA
132	20	50897361	67	Inserzione di 11 bp	G>GGAATGTCAGCC	15	0	0	15	100
134	20	62690925	87	Delezione di 16 bp	TCCTGGCTGGCCTGTGG>T	9	0	0	9	100
135	21	30300873	66	Inserzione di 11 bp	G>GATAAACTTTA	9	0	0	9	100
137	21	36710749	87	Delezione di 21 bp	ACTCAAGATAACTCATGTTATC>A	9	0	0	9	100
138	21	46644985	69	Delezione di 5 bp	GTTGTT>G	13	0	0	13	100
140	22	25750814	100	Inserzione di 6 bp	C>CAGGGCA	13	0	0	13	100
142	22	37409885	97	Inserzione di 5 bp	C>CTGTTT	13	0	0	13	100
144	22	47081407	92	Delezione di 10 bp	GGGCACAGGCA>G	7	0	0	7	100

## Studio 2

Nello studio sono stati utilizzati campioni di tessuto di tumore del colon-retto FFPE, proveniente da banche, e un saggio rappresentativo di due geni, confrontati con il metodo di riferimento, il sequenziamento bidirezionale Sanger (Sanger). Dei 1.183 soggetti totali, 441 soggetti hanno fornito risultati validi dal sequenziamento Sanger e dal saggio rappresentativo. Nella valutazione a livello di soggetto (Tabella 15), 230 dei 441 soggetti risultavano positivi secondo il sequenziamento Sanger (mutazione rilevata con il metodo Sanger). di questi, 227 erano positivi con il saggio rappresentativo. I restanti 211 di 441 soggetti erano negativi con il sequenziamento Sanger (nessuna mutazione rilevata da Sanger); di questi, 206 erano negativi con il saggio rappresentativo. Ciò corrisponde a una percentuale di concordanza positiva (positive percent agreement, PPA) del 98,7% e a una percentuale di concordanza negativa (negative percent agreement, NPA) del 97,6% (Tabella 15).

Tabella 15 Percentuale di concordanza positiva e negativa a livello di soggetto

Saggio rappresentativo	Sequenziamento Sanger		Totale
	Positivo	Negativo	
Positivo	227 <sup>1</sup>	5	232
Negativo	3 <sup>2</sup>	206	209
Performance (Prestazioni) Riepilogo			
Concordanza Statistica	Punto stimato		IC esatto al 95%

Saggio rappresentativo	Sequenziamento Sanger		Totale
	Positivo	Negativo	
Totale	230	211	441

Performance (Prestazioni) Riepilogo		
Concordanza Statistica	Punto stimato	IC esatto al 95%
PPA	227/230 = 98,7%	[96,2%, 99,7%]
NPA	206/211 = 97,6%	[94,6%, 99,2%]

<sup>1</sup>Sono state rilevate 224 corrispondenze esatte per i risultati all'interno dello stesso soggetto, a livello di tutte le mutazioni. Per due soggetti, MiSeqDx ha rilevato la mutazione rilevata da Sanger e un'ulteriore mutazione. Per un soggetto, MiSeqDx e Sanger hanno rilevato mutazioni diverse.

<sup>2</sup>Un soggetto aveva due mutazioni rilevate da Sanger. Due soggetti avevano una mutazione rilevata da Sanger.

### Studio 3

Questo studio ha valutato le librerie di DNA preparate con campioni FFPE su diversi tipi di campioni. Un totale di 109 campioni FFPE da otto tessuti diversi (colon, ovaio, pancreas, surrene, vescica, fegato, tiroide e mammella) con almeno 11 campioni FFPE a rappresentare ciascun tipo di tessuto. Il tessuto surrenale includeva metastasi da tumore all'esofago, ai polmoni e al colon. Gli altri tessuti presentavano tumori primari. Questo studio ha utilizzato un saggio rappresentativo progettato per interrogare 26 geni che coprono 21.577 basi su 17 cromosomi diversi. È stato sequenziato un totale di sei geni diversi (*KRAS*, *NRAS*, *TP53*, *PIK3CA*, *EGFR* e *BRAF*) con il metodo Sanger; ciascun tessuto tumorale presentava 1-3 geni sequenziati con metodo Sanger in base alla prevalenza prevista delle mutazioni somatiche di quel tumore. Il sequenziamento Sanger ha identificato 39 mutazioni somatiche SNV in 33 dei 109 campioni FFPE. Lo strumento MiSeqDx ha identificato 36 mutazioni somatiche SNV in 32 dei 109 campioni FFPE, con un falso negativo e due identificazioni di posizioni della variante non riuscite. La PPA era del 97,3%. Lo strumento MiSeqDx ha identificato 78.975 basi di riferimento nei 109 campioni FFPE, con 29 falsi positivi rispetto al sequenziamento Sanger e 3416 identificazioni non riuscite. La percentuale di concordanza negativa (Negative Percent Agreement, NPA) è risultata del 99,9%. I due metodi hanno evidenziato una concordanza riguardo alla delezione di due basi. La [Tabella 16](#) contiene un riepilogo dei risultati in base al tipo di tessuto.

Tabella 16 Percentuale di concordanza positiva e negativa in base al tipo di tessuto

Tipo di tessuto	N. di campioni	N. totale di varianti	N. totale di varianti TP	N. totale di varianti FN	N. totale di identificazioni TN	N. totale di identificazioni FP	N. totale di identificazioni non riuscite	PPA	NPA
Surrene	16	6	4	1	11.823	2	607	80	>99,9
Vescica	12	4	4	0	7.070	3	273	100	>99,9
Seno	16	3	3	0	13.439	7	479	100	99,9
Colon	11	6	5	0	8.720	2	133	100	>99,9
Fegato	13	3	3	0	7.984	1	59	100	>99,9
Ovaio	13	7	7	0	10.581	1	724	100	>99,99
Pancreas	17	7	7	0	11.929	12	489	100	99,9
Tiroide	11	3	3	0	7.429	1	652	100	>99,9
Totale	109	39	36	1	78975	29	3.416	97,3	>99,9

## Riproducibilità

Sono stati condotti due studi per valutare la riproducibilità dello strumento MiSeqDx con DNA estratto da campioni FFPE. Lo Studio 1 ha utilizzato più strumenti. Lo Studio 2 è stato svolto in più siti.

### Studio 1

La riproducibilità dello strumento MiSeqDx è stata determinata utilizzando due strumenti e due operatori addestrati, per un totale di otto corse. Il saggio rappresentativo, il contesto genomico dell'amplicone, i campioni e il metodo di riferimento sono gli stessi di quelli sopra descritti per lo Studio 1 di accuratezza. I risultati sono presentati in base ai singoli ampliconi per ciascuno strumento ([Tabella 17](#)), per dimostrare la riproducibilità delle identificazioni tra strumenti. La percentuale di identificazioni corrette include le identificazioni errate e le identificazioni non riuscite (l'identificazione di varianti non soddisfa uno o più filtri). Gli strumenti hanno generato numeri simili di identificazioni non riuscite dovute al determinato amplicone. Per lo strumento MiSeqDx 1 è stata osservata una singola identificazione errata all'interno di una regione di confidenza, in base alla definizione dello standard di riferimento Platinum Genomes. L'identificazione errata corrispondeva a un'identificazione falsa positiva di una variante di inserzione nell'amplicone 64 che interrogava il cromosoma 10 nelle posizioni da 55892599 a 55892687. L'amplicone presentava una ripetizione dinucleotidica di 11.

Tabella 17 Risultati dello studio di riproducibilità tra strumenti per lo strumento MiSeqDx (a livello di amplicone)

Amplicone	Cromosoma	Dimensione del frammento analizzato	Basi nella regione affidabile	Contenuto genomico dell'amplicone	Contenuto in GC	MiSeqDx 1			MiSeqDx 2		
						Totale di identificazioni corrette	Totale di identificazioni errate	N. totale di identificazioni non riuscite	Totale di identificazioni corrette	Totale di identificazioni errate	N. totale di identificazioni non riuscite
1	1	93	93	Indel	0,22	5.580	0	0	5.580	0	0
2	1	79	79	PolyA (5), PolyC (5), indel	0,38	4.740	0	0	4.740	0	0
3	1	91	91	Indel	0,4	5.448	0	12	5.453	0	8
4	1	92	92	Indel	0,49	5518	0	2	5518	0	2
5	1	81	81	PolyG (5)	0,69	4.858	0	2	4.860	0	0
6	1	70	70	PolyT (10), indel	0,39	4.200	0	0	4.200	0	0
7	1	88	88	PolyA (5), CT (3), TAA(3), indel	0,27	5.279	0	1	5.279	0	1
8	2	90	90	Indel	0,28	5.400	0	0	5.400	0	0
9	2	80	80	Indel	0,38	4.800	0	0	4.800	0	0
10	2	85	81	N/D	0,65	4.859	0	1	4.859	0	1
11	2	75	75	PolyA (8)	0,35	4.468	0	40	4.468	0	40
12	2	88	88	PolyT (5)	0,42	5.280	0	0	5.280	0	0
13	2	87	87	PolyT (5), indel	0,31	5.211	0	43	5.214	0	40
14	2	91	91	PolyT (5), PolyA (6), indel	0,3	5.453	0	7	5.449	0	11
15	2	93	93	N/D	0,43	5.579	0	1	5.579	0	1
16	2	73	73	PolyT (5), indel	0,42	4.378	0	2	4.379	0	1
17	3	93	93	AT(3), indel	0,27	5.396	0	184	5.396	0	184
18	3	83	83	N/D	0,43	4.980	0	0	4.980	0	0



Amplicone	Cromosoma	Dimensione del frammento analizzato	Basi nella regione affidabile	Contenuto genomico dell'amplicone	Contenuto in GC	MiSeqDx 1			MiSeqDx 2		
						Totale di identificazioni corrette	Totale di identificazioni errate	N. totale di identificazioni non riuscite	Totale di identificazioni corrette	Totale di identificazioni errate	N. totale di identificazioni non riuscite
19	3	70	70	CT(3), indel	0,49	4193	0	7	4194	0	6
20	3	88	88	PolyA (5), PolyT (5), PolyA (9), TG (3)	0,41	5.220	0	120	5.220	0	120
21	3	75	74	Indel	0,57	4.432	0	8	4.432	0	8
22	4	78	78	PolyA (6)	0,26	4.676	0	4	4.676	0	4
23	4	97	97	PolyG (6), PolyT (5), PolyA (5)	0,42	5.820	0	0	5.820	0	0
24	4	78	78	N/D	0,29	4.679	0	1	4.677	0	3
25	4	66	62	PolyA (5), indel	0,36	3.720	0	0	3.720	0	0
26	4	71	69	PolyA (5)	0,46	4.140	0	0	4.140	0	0
27	4	62	62	PolyA (7), indel	0,27	3.676	0	45	3.671	0	51
28	5	78	75	N/D	0,78	3.368	0	1.132	3.485	0	1.015
29	5	84	84	GT(3), CCA (3)	0,62	5.040	0	0	5.040	0	0
30	5	64	64	N/D	0,39	3.840	0	0	3.840	0	0
31	5	83	83	PolyA (6), indel	0,3	4.979	0	1	4.980	0	0
32	5	67	67	PolyT (5)	0,37	4.020	0	0	4.020	0	0
33	5	91	91	CT(4), AG(3)	0,55	5.460	0	0	5.460	0	0
34	5	75	75	Indel	0,43	4.498	0	6	4.500	0	1
35	6	102	102	PolyG (6)	0,68	6.120	0	0	6.120	0	0
36	6	92	92	Indel	0,63	5520	0	0	5520	0	0
37	6	100	94	GCT(5), indel	0,61	5.532	0	108	5.532	0	108

Amplicone	Cromosoma	Dimensione del frammento analizzato	Basi nella regione affidabile	Contenuto genomico dell'amplicone	Contenuto in GC	MiSeqDx 1			MiSeqDx 2		
						Totale di identificazioni corrette	Totale di identificazioni errate	N. totale di identificazioni non riuscite	Totale di identificazioni corrette	Totale di identificazioni errate	N. totale di identificazioni non riuscite
38	6	98	98	Poly T (5), TCT(3), CTT (3)	0,55	5.820	0	60	5.820	0	60
39	6	95	95	Indel	0,53	5.697	0	3	5.698	0	2
40	6	86	86	PolyC (6)	0,7	5159	0	1	5.160	0	0
41	6	95	94	PolyG (5), indel	0,61	5.638	0	2	5.638	0	2
42	6	91	91	PolyA (5)	0,44	5.460	0	0	5.460	0	0
43	7	73	73	N/D	0,44	4.380	0	0	4.380	0	0
44	7	88	88	Indel	0,35	5.279	0	1	5.276	0	4
45	7	87	87	PolyA (7), AG (4)	0,26	5.184	0	36	5.181	0	39
46	7	85	85	Indel	0,38	5.100	0	0	5.100	0	0
47	7	90	90	PolyG (5), indel	0,62	5.398	0	2	5.399	0	1
48	7	91	91	PolyG (6), PolyC (6), indel	0,71	5.460	0	0	5.459	0	1
49	7	81	66	N/D	0,31	3.960	0	0	3.960	0	0
50	7	93	93	Indel	0,35	5.580	0	0	5.579	0	1
51	8	83	83	N/D	0,42	4.980	0	0	4.980	0	0
52	8	90	89	PolyG (7), CTC(4), indel	0,61	5.219	0	121	5.220	0	120
53	8	67	67	Indel	0,3	4.020	0	0	4.020	0	0
54	9	98	98	PolyG (6)	0,67	5879	0	1	5.880	0	0
55	9	65	65	Indel	0,32	3.894	0	6	3.895	0	5
56	9	96	96	N/D	0,49	5.760	0	0	5.760	0	0
57	9	83	83	AT(3)	0,37	4.973	0	7	4.978	0	2

Amplicone	Cromosoma	Dimensione del frammento analizzato	Basi nella regione affidabile	Contenuto genomico dell'amplicone	Contenuto in GC	MiSeqDx 1			MiSeqDx 2		
						Totale di identificazioni corrette	Totale di identificazioni errate	N. totale di identificazioni non riuscite	Totale di identificazioni corrette	Totale di identificazioni errate	N. totale di identificazioni non riuscite
58	9	97	97	PolyC (6), indel	0,68	5.817	0	3	5.818	0	2
59	10	79	78	PolyG (5), indel	0,47	4.679	0	1	4.680	0	0
60	10	98	91	GC(3)	0,87	450	0	5.010	632	0	4.828
61	10	79	79	PolyT (5)	0,3	4.740	0	0	4.740	0	0
62	10	90	90	PolyA (5), PolyT (5)	0,2	5.400	0	0	5.400	0	0
63	10	95	95	Indel	0,35	5.699	0	1	5.699	0	1
64	10	89	88	AC(11), indel	0,42	5.157	0	276	5.153	2	273
65	10	80	80	N/D	0,49	4.800	0	0	4.800	0	0
66	10	81	81	N/D	0,51	4.860	0	0	4.860	0	0
67	11	97	96	N/D	0,45	5.760	0	0	5.760	0	0
68	11	70	70	Indel	0,46	4.199	0	2	4.200	0	1
69	11	100	100	N/D	0,65	5.999	0	1	5.998	0	2
70	11	62	62	Indel	0,37	3.720	0	0	3.720	0	0
71	11	102	102	N/D	0,59	6.120	0	0	6.118	0	2
72	11	73	73	PolyA (5)	0,4	4.380	0	0	4.380	0	0
73	11	85	85	N/D	0,42	5.100	0	0	5.100	0	0
74	11	91	91	PolyG (6)	0,55	5.437	0	23	5.441	0	19
75	11	85	85	Indel	0,53	5.100	0	0	5.100	0	0
76	11	85	85	Poly A (5), CA (3), indel	0,34	5.100	0	0	5.100	0	0
77	11	85	85	GA(3)	0,52	5.100	0	0	5.100	0	0
78	12	84	84	PolyC (5), indel	0,52	5.040	0	60	5.038	0	63

Amplicone	Cromosoma	Dimensione del frammento analizzato	Basi nella regione affidabile	Contenuto genomico dell'amplicone	Contenuto in GC	MiSeqDx 1			MiSeqDx 2		
						Totale di identificazioni corrette	Totale di identificazioni errate	N. totale di identificazioni non riuscite	Totale di identificazioni corrette	Totale di identificazioni errate	N. totale di identificazioni non riuscite
79	12	93	93	PolyA (7), AC (4)	0,33	5.577	0	3	5573	0	7
80	12	81	81	N/D	0,49	4.860	0	0	4.860	0	0
81	12	71	71	PolyA (6)	0,35	4.260	0	0	4.260	0	0
82	2	95	95	PolyG (5)	0,68	5605	0	95	5605	0	95
83	13	73	73	N/D	0,52	4.380	0	0	4.379	0	1
84	13	89	88	PolyA (5), PolyT (7), PolyA (7), indel	0,22	5.220	0	60	5.220	0	60
85	13	89	89	Indel	0,49	5.340	0	0	5.340	0	0
86	13	77	77	Indel	0,39	4.620	0	0	4.620	0	0
87	14	67	67	GA(3), TA(3)	0,39	4.020	0	0	4.020	0	0
88	14	83	83	N/D	0,25	4.980	0	0	4.980	0	0
89	14	73	72	PolyT (5), indel	0,19	4.173	0	147	4.173	0	147
90	14	81	81	Indel	0,38	4.860	0	2	4.860	0	0
91	14	91	91	Indel	0,35	5.459	0	1	5.460	0	0
92	14	66	66	PolyA (5)	0,41	3.900	0	240	3.900	0	240
93	14	94	94	Indel	0,62	5.637	0	3	5.637	0	3
94	15	98	96	PolyC (5)	0,45	5.760	0	0	5.760	0	0
95	15	68	68	Indel	0,25	4.079	0	1	4.078	0	2
96	15	95	93	PolyG (5), indel	0,68	5.475	0	105	5.487	0	93
97	15	95	95	PolyT (6)	0,43	5.699	0	1	5.700	0	0
98	15	72	71	Indel	0,65	4.260	0	0	4.260	0	0
99	15	91	91	N/D	0,36	5.460	0	0	5.460	0	0

Amplicone	Cromosoma	Dimensione del frammento analizzato	Basi nella regione affidabile	Contenuto genomico dell'amplicone	Contenuto in GC	MiSeqDx 1			MiSeqDx 2		
						Totale di identificazioni corrette	Totale di identificazioni errate	N. totale di identificazioni non riuscite	Totale di identificazioni corrette	Totale di identificazioni errate	N. totale di identificazioni non riuscite
100	15	70	70	Indel	0,56	4.200	0	0	4.200	0	0
101	16	63	63	N/D	0,27	3.780	0	0	780	0	0
102	16	95	95	PolyC (5)	0,67	5.700	0	0	5.700	0	0
103	16	87	87	TA(3)	0,41	5.220	0	0	5.220	0	0
104	16	104	104	PolyC (5)	0,67	6.238	0	3	6.238	0	3
105	16	91	91	PolyT (5), indel	0,37	5.443	0	17	5.444	0	16
106	17	89	89	GC(3)	0,64	5.251	0	89	5.339	0	1
107	17	87	87	PolyC (5), indel	0,67	5.212	0	8	5.212	0	8
108	17	91	91	Indel	0,46	5.459	0	1	5.459	0	1
109	17	93	93	Indel	0,26	5.580	0	0	5.580	0	0
110	17	91	89	PolyT (5)	0,54	5.340	0	0	5.340	0	0
111	17	84	82	Poly A (13), indel (x2)	0,29	4.860	0	308	4.860	0	07
112	17	91	91	PolyA (5)	0,34	5.459	0	1	5.459	0	1
113	17	92	92	PolyA (7), AT (3), AT(4), AT (4), indel	0,26	5.460	0	60	5.460	0	60
114	17	79	79	Indel	0,41	4.699	0	41	4.700	0	40
115	17	86	86	PolyT (7)	0,22	5.153	0	7	5156	0	4
116	17	84	84	GAG(3)	0,62	5.039	0	1	5.039	0	1
117	18	67	67	GA(3)	0,31	4.020	0	0	4.020	0	0
118	18	91	91	N/D	0,37	5.460	0	0	5.460	0	0
119	18	69	69	PolyA (6), TG (3)	0,43	4.132	0	8	4.131	0	9

Amplicone	Cromosoma	Dimensione del frammento analizzato	Basi nella regione affidabile	Contenuto genomico dell'amplicone	Contenuto in GC	MiSeqDx 1			MiSeqDx 2		
						Totale di identificazioni corrette	Totale di identificazioni errate	N. totale di identificazioni non riuscite	Totale di identificazioni corrette	Totale di identificazioni errate	N. totale di identificazioni non riuscite
120	18	75	75	PolyA (5), indel	0,37	4.475	0	85	4.480	0	79
121	18	81	81	CTC(3), indel	0,47	4.860	0	0	4.860	0	0
122	18	85	85	CT(3), indel	0,45	5.098	0	2	5.098	0	2
123	18	65	65	N/D	0,48	3.900	0	0	3.900	0	0
124	19	99	99	N/D	0,59	5.926	0	14	5.924	0	16
125	19	74	74	N/D	0,68	4.440	0	0	4.438	0	2
126	19	70	70	N/D	0,64	4.199	0	1	4.200	0	0
127	20	94	94	N/D	0,61	5.640	0	1	5.638	0	3
128	20	82	82	AC(3)	0,59	4.920	0	0	4.920	0	0
129	20	76	76	CT(3)	0,58	4.559	0	1	4.558	0	2
130	20	70	70	GT(3), TG(4), indel	0,46	4.200	0	0	4.200	0	0
131	20	101	101	Indel	0,63	6.060	0	0	6.060	0	0
132	20	67	67	Indel	0,36	4.020	0	31	4.020	0	25
133	20	91	88	PolyG (6)	0,73	5.277	0	3	5.274	0	6
134	20	87	87	Indel	0,57	5.218	0	2	5.218	0	2
135	21	66	66	Indel	0,35	3.959	0	1	3.957	0	3
136	21	98	98	PolyT (6), CA (3)	0,54	5.880	0	0	5.880	0	0
137	21	87	87	GT(3), indel	0,39	5.220	0	0	5.220	0	0
138	21	69	69	PolyA (6), AG (3), indel	0,32	4.119	0	31	4.113	0	37
139	21	90	90	PolyT (5), PolyA (6)	0,5	5.399	0	1	5.399	0	1
140	22	100	100	Indel	0,63	5.998	0	7	5.997	0	5
141	22	97	97	N/D	0,68	5.819	0	1	5.819	0	1

Amplicone	Cromosoma	Dimensione del frammento analizzato	Basi nella regione affidabile	Contenuto genomico dell'amplicone	Contenuto in GC	MiSeqDx 1			MiSeqDx 2		
						Totale di identificazioni corrette	Totale di identificazioni errate	N. totale di identificazioni non riuscite	Totale di identificazioni corrette	Totale di identificazioni errate	N. totale di identificazioni non riuscite
142	22	97	97	Indel	0,46	5.818	0	2	5.816	0	4
143	22	99	99	N/D	0,6	5.940	0	0	5.940	0	0
144	22	92	92	Indel	0,66	5.519	0	1	5.519	0	1
145	X	69	69	PolyT (5)	0,26	4.139	0	1	4.140	0	0
146	X	69	69	PolyC (5)	0,62	4.136	0	4	4.137	0	3
147	X	71	71	N/D	0,52	4.260	0	0	4.260	0	0
148	Y	65	0	N/D	0,55	0	0	0	0	0	0
149	Y	91	0	N/D	0,48	0	0	0	0	0	0
150	Y	71	0	PolyA (5)	0,37	0	0	0	0	0	0

I risultati dello studio di riproducibilità sono stati analizzati in base all'operatore utilizzando la frequenza di varianti (Tabella 18). Questa analisi ha dimostrato che, per gli operatori, le frequenze delle varianti erano coerenti. Sono presentate le frequenze delle varianti medie alla deviazione standard di +/- 1.

Tabella 18 Risultati da operatore a operatore per lo strumento MiSeqDx

Intervallo delle frequenze delle varianti	N. varianti univoche	N. totale di varianti analizzate dall'operatore 1	N. totale di varianti analizzate dall'operatore 2	Frequenza della variante media (DS) riportata dall'operatore 1	Frequenza della variante media (DS) riportata dall'operatore 2
Frequenza elevata (circa 100%)	1.112	1.072	1.072	0,96 +/- 0,05	0,96 +/- 0,05
Frequenza media (circa 50%)	3.240	3.151	3.161	0,49 +/- 0,04	0,49 +/- 0,04
Frequenza bassa (3-7%)	620	618	612	0,05 +/- 0,01	0,05 +/- 0,01

I risultati dello studio di riproducibilità per ogni campione sono riportati combinati per tutte le otto corse (Tabella 19). Il rilevamento è stato valutato separatamente per ciascun tipo di variante: SNV, inserzioni e delezioni separatamente. Sono state escluse le posizioni di riferimento. Questa analisi dimostra che i risultati per le varianti erano riproducibili sui campioni.

Tabella 19 Concordezza dei risultati di identificazione delle basi dello strumento MiSeqDx per campione

Campione	SNV				Inserzioni				Delezioni			
	N. totale	N. totale TP	N. totale FP	N. totale FN	N. totale	N. totale TP	N. totale FP	N. totale FN	N. totale	N. totale TP	N. totale FP	N. totale FN
GM12877	592	574	2	0	336	336	0	0	228	272	0	0
GM12878	1.456	1.432	0	0	320	304	0	0	384	352	0	0
GM12879	912	896	0	0	336	320	0	0	288	272	0	0
GM12885	1.200	1.192	0	0	400	384	0	0	352	320	0	0
GM12886	1.104	1.104	0	0	368	352	0	0	368	352	0	0
GM12877-D1 <sup>1</sup>	3.640	3.582	0	0	800	760	0	0	960	880	0	0
GM12877-D2 <sup>2</sup>	400	398	0	0	520	516	0	0	560	556	0	0

<sup>1</sup> Varianti con una frequenza superiore al 20%.

<sup>2</sup> Varianti con una frequenza inferiore al 20%.

I dati forniti dalle otto corse di questo studio di riproducibilità dimostrano che lo strumento MiSeqDx è in grado di sequenziare in modo coerente:

- contenuto in GC  $\geq$  19% (tutte le basi identificate in 120 dei 120 ampliconi sequenziati con 19% di contenuto in GC identificato correttamente con una percentuale di rilevamento non riuscito del 3,4%)
- Contenuto in GC  $\leq$  73% (tutte le basi identificate in 120 dei 120 ampliconi sequenziati con il 73% di contenuto in GC identificato correttamente con una percentuale di rilevamento non riuscito dello 0,1%)
- Lunghezze PolyA  $\leq$  8 (la ripetizione PolyA di otto nucleotidi era stata identificata correttamente in 120 dei 120 ampliconi sequenziati contenenti PolyA = 8)



- Lunghezze PolyT  $\leq 10$  (la ripetizione PolyT di 10 nucleotidi era stata identificata correttamente in 120 dei 120 ampliconi sequenziati contenenti PolyT = 10)
- Lunghezze PolyG  $\leq 6$  (la ripetizione PolyG di sei nucleotidi era stata identificata correttamente in 720 dei 720 ampliconi sequenziati contenenti PolyG = 6)
- Lunghezze PolyC  $\leq 6$  (la ripetizione PolyC di sei nucleotidi era stata identificata correttamente in 359 dei 360 ampliconi sequenziati contenenti PolyC = 6 con un'identificazione non riuscita)
- Lunghezze di dinucleotidi ripetuti  $\leq 4x$  (tutte le basi identificate in 600 dei 600 ampliconi sequenziati con 4x di ripetizione di dinucleotidi sono state identificate correttamente con una percentuale di rilevamento non riuscito dello 0,4%)
- Lunghezze di trinucleotidi ripetuti  $\leq 5x$  (tutte le basi identificate in 120 dei 120 ampliconi sequenziati con 5x di ripetizione di trinucleotide sono state identificate correttamente con una percentuale di rilevamento non riuscito dell'1,9%)
- 24 o meno inserzioni delle basi e 25 o meno delezioni della basi
  - 24 inserzioni delle basi identificate correttamente in 120 dei 120 campioni
  - 25 delezioni delle basi identificate correttamente in 182 campioni e riportate come identificazione non riuscita in due dei 184 campioni

## Studio 2

È stato condotto uno studio esterno per valutare la riproducibilità del saggio rappresentativo di due geni, descritto nello Studio 2 sull'accuratezza, in tre siti di analisi esterni (due operatori per sito), con un lotto di reagenti e tre giorni di analisi non consecutivi. L'analisi è stata condotta con sei pannelli di campioni ben caratterizzati di campioni di DNA genomico ottenuti da campioni clinici o linee cellulari FFPE. Ciascun pannello consisteva di 10 membri, per un totale di 60 membri nei pannelli.

I 60 membri del pannello consistevano di duplicati di quattro campioni wild type univoci (per le mutazioni del pannello), 12 campioni di mutazioni univoche (con una singola mutazione) preparati sia al livello alto che al livello basso della frequenza della mutazione e due campioni di mutazione univoca (con una sola mutazione) preparati solo al livello basso della frequenza della mutazione. Ciascun campione univoco a livello di frequenza campione/mutazione (analizzato in duplicati in ogni corsa) presentava 36 possibili risultati (due replicati  $\times$  due operatori  $\times$  tre giorni  $\times$  3 siti) se tutti i risultati erano validi.

L'identificazione prevista espressa in percentuale (PEC) su tutte le varianti positive e negative è stata valutata confrontando il risultato del saggio rappresentativo con il risultato della mutazione prevista (mutazione prevista rilevata o non rilevata) in ciascun campione. La PEC è calcolata come 100% delle volte il numero di identificazioni previste per il numero di identificazioni tentate. L'intervallo di confidenza bilaterale del 95% è calcolato utilizzando il metodo del punteggio Wilson.

Combinando i siti, le percentuali di superamento dei campioni erano  $\geq 94,7\%$  per la prima corsa del campione o nei campioni analizzati nelle corse che erano validi al primo passaggio "first pass". La PEC a livello di mutazione su tutti i campioni contenenti mutazioni era di 99,6% (905/909) (IC al 95%; 98,9, 99,8). Il numero delle identificazioni tentate su tutte le 56 mutazioni del pannello (indipendentemente dal fatto che una mutazione rilevata fosse prevista o no) per tutti i campioni validi era 58.856 ( $56 \times 1.051$ ). Tra queste 58.856 osservazioni a livello di mutazione, sono stati rilevati solo sei casi in cui i risultati osservati e quelli previsti erano discordanti. La PEC a livello di mutazione su tutte le varianti positive e negative di tutti i membri del pannello mutanti e wild type combinati è stata del 99,99% (58.850/58.856).

### Sensibilità analitica: limite del bianco (Limit of Blank, LoB) e limite del rilevamento (Limit of Detection, LoD)

Questo studio ha verificato il valore di cutoff del saggio e determinato il limite del rilevamento (LoD) per MiSeqDx con un pannello rappresentativo. In breve, le linee cellulari ben caratterizzate GM12878 e GM12877 Platinum Genome erano campioni fissati in formalina e inclusi in paraffina dai quali è stato estratto il DNA. GM12878 è stato diluito con GM12877 in modo tale che le frequenze delle varianti di settanta varianti (52 SNV, nove inserzioni e nove delezioni) fossero vicine allo 0,05. I due campioni di DNA sono stati analizzati da due operatori, utilizzando due strumenti e due lotti di reagenti, per un totale di 10 corse di sequenziamento MiSeqDx. In questo modo sono stati ottenuti 40 replicati per ciascuna variante in GM12878 e 60 replicati per ciascuna coordinata wild type corrispondente in GM12877 per ciascun lotto di reagenti. LoB e LoD sono stati calcolati con l'approccio classico indicato da CLSI EP17-A2 utilizzando l'opzione non parametrica. LoB e LoD sono stati calcolati per SNV, inserzioni e delezioni separatamente raggruppando in pool le frequenze delle varianti per un dato tipo di variante. L'errore di tipo I era definito come 0,01 e l'errore di tipo II era definito come 0,05.

Per LoB, le frequenze delle varianti raggruppate in pool sono state ordinate dalla più bassa alla più alta ed è stata calcolata la posizione 99 per ciascun lotto di reagenti per ciascun tipo di variante (Tabella 20). Il software MiSeqDx utilizza un valore di cutoff (l'effettivo LoB) di 0,026 di frequenza della variante per determinare il rilevamento qualitativo delle varianti. I limiti calcolati del campione bianco hanno verificato che questo cutoff ha comportato un errore di tipo I di non più di 0,01.

Tabella 20 Limite del bianco

Tipo di variante	Numero totale delle frequenze delle varianti	LoB del lotto di reagenti 1 (%)	LoB del lotto di reagenti 2 (%)
SNV	3.120	0,87	0,75

Tipo di variante	Numero totale delle frequenze delle varianti	LoB del lotto di reagenti 1 (%)	LoB del lotto di reagenti 2 (%)
Inserzione	540	0,79	0,60
Delezione	540	0,96	0,84

Per il LoD è stata calcolata la percentuale della frequenza delle singole mutazioni per ciascun lotto di reagenti per ciascun tipo di variante che risultava al di sotto del valore soglia di 0,026 (Tabella 21). Poiché le percentuali erano inferiori all'errore di tipo II del 5% (0,05), come valore LoD è stata calcolata la mediana delle frequenze delle varianti combinate (Tabella 21). È stato preso come LoD per ciascun tipo di variante il valore maggiore dei due valori calcolati per i due lotti di reagenti: 5,45% per le SNV, 4,88% per le inserzioni e 5,44% per le delezioni.

Tabella 21 Limite del rilevamento

Lotto di reagenti	Tipo di variante	Numero totale delle frequenze delle varianti	N. di misurazioni VF < 2,6%	% di misurazioni VF < 2,6%	Limite del rilevamento (%)
1	SNV	2.080	4	0,20	5,45
	Inserzione	360	0	0,00	4,86
	Delezione	360	2	0,60	5,44
2	SNV	2.080	26	1,30	5,44
	Inserzione	360	0	0,00	4,88
	Delezione	360	0	0,00	5,24

I seguenti studi dimostrano le caratteristiche prestazionali dello strumento MiSeqDx con un altro saggio rappresentativo mirato a 56 mutazioni in due oncogeni clinicamente rilevanti (pannello delle mutazioni). Il pannello delle mutazioni è stato progettato per rilevare specificatamente le 56 mutazioni in due geni del cancro rilevanti dal punto di vista clinico (Gene 1 e Gene 2). Il saggio determina contemporaneamente la presenza o l'assenza di ciascuna delle 56 mutazioni in ogni campione sequenziato. Il metodo di riferimento per questi studi è stato il sequenziamento bidirezionale Sanger.

## Precisione da lotto a lotto

È stato condotto uno studio di precisione da lotto a lotto per valutare le prestazioni dello strumento MiSeqDx su lotti di kit di reagenti fabbricati (che consistono di qualificazione del campione, preparazione delle librerie e reagenti per il sequenziamento) utilizzando un saggio rappresentativo di due geni con un pannello di cinque campioni FFPE misti che soddisfacevano i requisiti di qualificazione. Ciascun campione FFPE conteneva due mutazioni univoche: una al livello più basso di frequenza della mutazione (circa 8%) e una al livello più alto di frequenza della mutazione (circa 14%). Dodici (12) osservazioni di ciascuno dei cinque campioni misti sono stati raccolti in tre giorni non consecutivi con tre lotti di kit di reagenti. Il numero totale di osservazioni per lo studio su tutti i lotti di reagenti era di 180 osservazioni su tutti i campioni misti e 360 osservazioni su tutti i livelli di frequenza della mutazione. Su tutti i lotti e giorni, il 99,7% (359/360) delle osservazioni ha mostrato il risultato della mutazione prevista. Una mutazione a bassa frequenza è stata erroneamente identificata come wild type. È stata eseguita un'analisi delle componenti della varianza per ciascun livello di mutazioni/frequenza di mutazione per stimare la variabilità del sistema. La deviazione standard totale era compresa tra 0,011 e 0,029. Il componente del lotto di reagenti della deviazione standard totale era compresa tra 0 e 0,015.

## Cronologia revisioni

Documento n.	Data	Descrizione della modifica
Documento n. 200006218 v02	Ottobre 2023	Aggiornamento dei riferimenti di Local Run Manager per la versione 4. Aggiornamento dell'etichettatura allo scopo di aggiungere i dettagli dello sponsor australiano. Correzione delle discrepanze tra foglio illustrativo e rapporto dello studio di verifica dello strumento per MiSeqDx.
Documento n. 200006218 v01	Maggio 2022	Aggiunta del foglio illustrativo di MiSeqDx Reagent Kit v3 Micro CN in Apparecchiatura e materiali forniti, non necessario. Aggiunte informazioni sui test per MiSeqDx Reagent Kit v3 Micro in Caratteristiche delle prestazioni. Rimozione nota di attenzione specifica per gli Stati Uniti da Avvertenze e precauzioni.
Documento n. 200006218 v00	Novembre 2021	Versione iniziale a supporto di MOS v4.0 e Local Run Manager v3.0.

## Brevetti e marchi di fabbrica

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 similari di alcun genere.

Al fine di garantire un uso sicuro e corretto dei prodotti qui descritti, le istruzioni riportate nel presente documento devono essere scrupolosamente ed esplicitamente seguite da personale qualificato e adeguatamente formato. 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 QUI CONTENUTE POSSONO CAUSARE DANNI AL/I PRODOTTO/I, LESIONI PERSONALI A UTENTI E TERZI E DANNI MATERIALI E RENDERANNO NULLA QUALSIASI GARANZIA APPLICABILE AL/I PRODOTTO/I.

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

© 2023 Illumina, Inc. Tutti i diritti riservati.

Tutti i marchi di fabbrica sono di proprietà di Illumina, Inc. o dei rispettivi proprietari. Per informazioni specifiche sui marchi di fabbrica, consultare la pagina web [www.illumina.com/company/legal.html](http://www.illumina.com/company/legal.html).

## Informazioni di contatto



Illumina, Inc.  
5200 Illumina Way  
San Diego, California 92122 U.S.A.  
+1.800.809.ILMN (4566)  
+1.858.202.4566 (fuori dal Nord America)  
techsupport@illumina.com  
www.illumina.com



Illumina Netherlands B.V.  
Steenoven 19  
5626 DK Eindhoven  
The Netherlands

### Sponsor australiano

Illumina Australia Pty Ltd  
Nursing Association Building  
Level 3, 535 Elizabeth Street  
Melbourne, VIC 3000  
Australia

## Etichettatura del prodotto

Per un riferimento completo dei simboli che si trovano sulla confezione del prodotto e sull'etichettatura, fare riferimento alla legenda dei simboli alla pagina [Web support.illumina.com](http://support.illumina.com) sulla scheda *Documentation* (Documentazione) per il kit in uso.