

# illumina® Bio-Rad® SureCell™ WTA 3' Library Prep Kit per il sistema ddSEQ™

Una panoramica approfondita dell'espressione genica grazie a una piattaforma di sequenziamento di nuova generazione sensibile, scalabile, dotata di elevata processività ed efficace in termini di costi

## Punti principali

- Caratterizzazione sensibile e imparziale di firme (signature) trascrizionali**  
 Rilevamento del gene altamente sensibile e riproducibile in singole cellule
- Flusso di lavoro completo per il sequenziamento dell'RNA di singole cellule**  
 Flusso di lavoro completamente supportato, sviluppato in collaborazione con gli innovatori nel campo della tecnologia
- Sequenziamento di nuova generazione altamente efficace integrato con una semplice analisi dei dati**  
 Sequenziamento comprovato Illumina abbinato a un software di analisi ottimizzato e di facile utilizzo

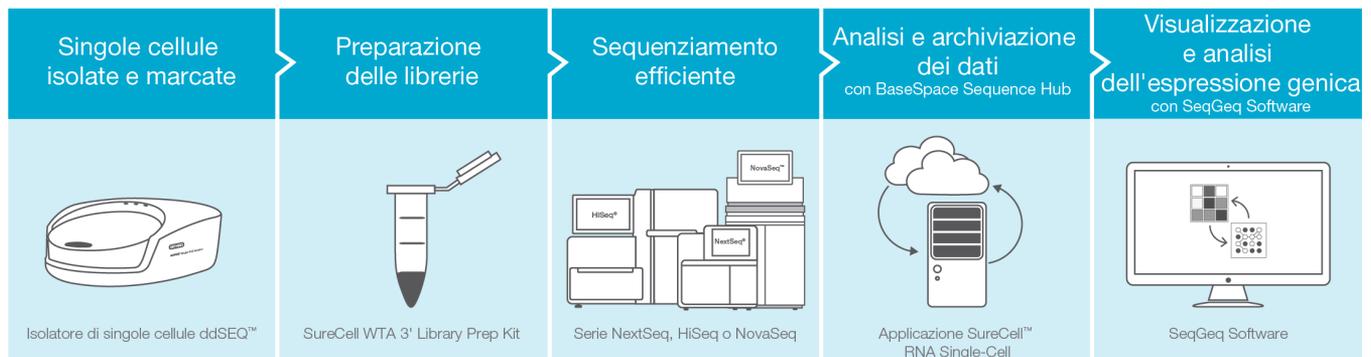
Per tenere fede alla promessa di una biologia mirata alla singola cellula, la soluzione Bio-Rad Single-Cell Sequencing Illumina combina l'innovativa tecnologia Bio-Rad Droplet Digital™(3) con le tecnologie NGS Illumina per la preparazione delle librerie, il sequenziamento e l'analisi. Questa nuova piattaforma fornisce un flusso di lavoro completo e di facile utilizzo per il sequenziamento dell'RNA di singole cellule (Figura 1), che permette di eseguire esperimenti controllati con più campioni, condizioni di trattamento e punti temporali. Creato e supportato grazie a una collaborazione tra i leader nella tecnologia, la soluzione Bio-Rad Single-Cell Sequencing Illumina permette di analizzare, in un solo esperimento, da centinaia a decine di migliaia di singole cellule per identificarne il trascrittoma. Questa metodologia NGS efficace e scalabile di preparazione dei campioni per singole cellule permette a più ricercatori di applicare la sensibilità e la precisione del sequenziamento dell'RNA a domande alle quali è possibile rispondere solo interrogando singole cellule.

## Introduzione

I complessi sistemi biologici sono determinati dalle funzioni coordinate delle singole cellule. Le tecnologie convenzionali che forniscono un ampio numero di dati del trascrittoma sono in grado di rivelare l'eterogeneità trascrizionale che è alla base di questa complessità (1). Il sequenziamento dell'RNA (RNA-Seq) di singole cellule permette l'analisi approfondita dell'espressione genica, che fornisce una panoramica dettagliata della funzione della cellula, della progressione della malattia e dell'efficacia terapeutica (2). Tuttavia, risulta ancora difficile generare in modo economico, semplice e ad elevata processività migliaia di librerie di singole cellule per il sequenziamento di nuova generazione (Next-Generation Sequencing, NGS) (2).

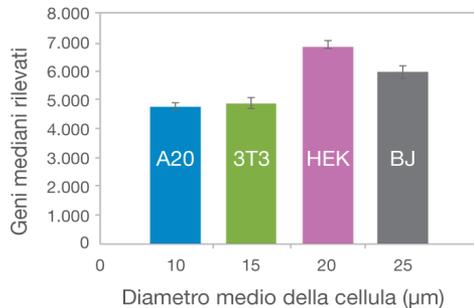
## Isolamento di singole cellule flessibile e scalabile

L'isolatore Bio-Rad ddSEQ Single-Cell Isolator (SCI) incapsula e separa le singole cellule in goccioline di dimensioni inferiori al nanolitro su una cartuccia monouso. Ciascuna cartuccia può contenere più campioni e più cartucce possono essere elaborate in parallelo, per permettere l'isolamento da centinaia a decine di migliaia di cellule al giorno. La lisi cellulare e la marcatura delle cellule si verificano all'interno delle singole goccioline, per permettere di tracciare le singole cellule durante tutto il flusso di lavoro. Questo metodo basato sulle goccioline non dipende dalla dimensione delle cellule di mammifero, pertanto fornisce un profilo imparziale di diverse popolazioni di cellule (Figura 2).



**Fig. 1. Flusso di lavoro del sequenziamento dell'RNA di singole cellule.** Il flusso di lavoro integra l'isolamento cellulare comprovato mediante Bio-Rad ddSEQ SCI, seguito dalla preparazione delle librerie mediante SureCell WTA 3' Library Prep Kit con la tecnologia Nextera, il sequenziamento Illumina e l'analisi dei dati mediante BaseSpace Sequence Hub e SeqGeq Analysis Software di FlowJo, LLC.

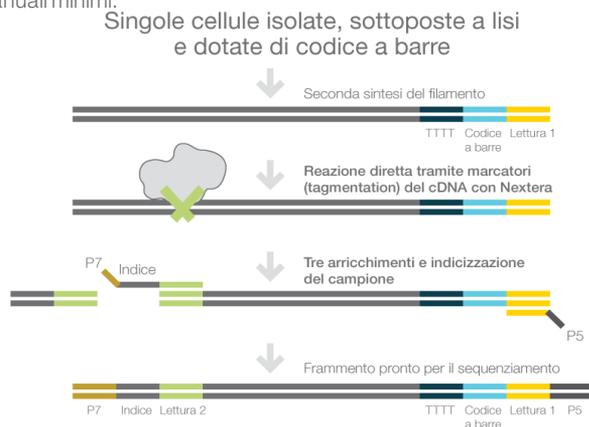
L'isolatore Bio-Rad ddSEQ SCI è in grado di elaborare i campioni in meno di cinque minuti. Eliminando lunghi flussi di lavoro sperimentali che possono alterare le firme (signature) trascrizionali, è possibile rilevare e tracciare risposte trascrizionali precise nel corso degli esperimenti.



**Fig. 2. Sequenziamento dell'RNA imparziale che non dipende dalla dimensione delle cellule.** Le linee cellulari murine (A20, NIH 3T3) e le linee cellulari umane (HEK, BJ) sono state elaborate con SureCell WTA 3' Library Prep Kit. Un numero coerentemente elevato di geni rilevati dimostra che il recupero di trascritti non è limitato dalla dimensione delle cellule.

## Chimica NGS avanzata per la preparazione delle librerie

Una sospensione di singola cellula viene caricata su una cartuccia ddSEQ M e le cellule vengono incapsulate e marcate dall'isolatore di singole cellule ddSEQ. La lisi e la marcatura vengono eseguite in ogni gocciolina. Le goccioline vengono separate e il cDNA viene raggruppato per la seconda sintesi del filamento. Le librerie vengono generate mediante reazione diretta tramite marcatori (tagmentation) del cDNA usando la tecnologia Nextera®, senza dover eseguire la frammentazione (shearing) o la preamplificazione (Figura 3). Alla reazione tramite marcatori (tagmentation) segue l'arricchimento dell'estremità 3' e l'indicizzazione del campione per preparare fino a 24 librerie indicizzate e pronte per il sequenziamento con interventi manuali minimi.



**Fig. 3. Descrizione generale di SureCell WTA 3' Library Prep Kit per il sistema ddSEQ.** Dopo l'isolamento, la lisi e la marcatura delle cellule, il cDNA viene raggruppato per la seconda sintesi del filamento. Alla generazione delle librerie mediante reazione diretta tramite marcatori (tagmentation) del cDNA segue l'arricchimento dell'estremità 3' e l'indicizzazione del campione.

## Sequenziamento affidabile con i sistemi NGS Illumina

Le librerie preparate di singole cellule vengono caricate direttamente su un sistema della serie MiSeq®, NextSeq®, HiSeq® o NovaSeq™ Illumina per il sequenziamento. Questi sistemi di sequenziamento utilizzano la chimica mediante sintesi (Sequencing By Synthesis, SBS) Illumina leader nel settore, utilizzata in più del 90% di tutti i sequenziamenti NGS in tutto il mondo.\* Questa tecnologia comune assicura la coerenza e l'affidabilità dei dati del sequenziamento su tutte le piattaforme NGS Illumina. I sistemi di sequenziamento delle serie MiSeq, NextSeq, HiSeq e NovaSeq offrono la flessibilità per un'ampia gamma di applicazioni e la scalabilità del rendimento per supportare diverse dimensioni di studi.

Il software di controllo dello strumento è di facile utilizzo e guida i ricercatori nelle procedure di caricamento dei campioni e di impostazione della corsa mediante un'interfaccia utente intuitiva. Inoltre permette il monitoraggio in tempo reale di corse di sequenziamento completamente automatizzate integrate sullo strumento o a distanza.

## Analisi semplificata dei dati in BaseSpace® Sequence Hub

I dati del sequenziamento della singola cellula possono essere trasferiti, archiviati e analizzati istantaneamente in modo sicuro in BaseSpace Sequence Hub, l'ambiente di calcolo genomico sul cloud Illumina. BaseSpace Sequence Hub fornisce un'ampia raccolta di applicazioni di BaseSpace. Una serie di esigenze di analisi dei dati comuni come allineamento, identificazione di varianti, ecc. è supportata da strumenti commerciali e open-source. Queste applicazioni dispongono di interfacce utente da utilizzare premendo solo un pulsante e sono progettate per essere utilizzate senza dover possedere esperienza bioinformatica.

L'applicazione SureCell RNA Single-Cell è stata progettata per supportare l'analisi dei dati per la soluzione Bio-Rad Single-Cell Sequencing Illumina. L'applicazione SureCell RNA Single-Cell permette l'analisi dei dati ottimizzata per un massimo di 96 campioni su più corse di sequenziamento e comprende:

- Pratiche metriche per il controllo qualità (Quality Control, QC) del sequenziamento
- Facile assegnazione di trascritti univoci su singole cellule
- Matrici e report dell'espressione genica esportabili e scaricabili
- Diverse opzioni di analisi per l'identificazione di sottopopolazioni e geni dell'espressione differenziale

## Visualizzazione avanzata dei dati con SeqGeq Software



SeqGeq Software è un'applicazione software per l'analisi avanzata dei dati, l'esplorazione e la visualizzazione dei dati dell'espressione genica della cellula singola, sviluppati da FlowJo, LLC per la soluzione Bio-Rad Single-Cell Sequencing Illumina. SeqGeq Software dispone di strumenti per un'efficace riduzione dei dati e l'identificazione della popolazione. L'integrazione diretta con BaseSpace Sequence Hub consente la visualizzazione e l'analisi dei dati di espressione con mappatura statistica a colori delle singole cellule, mappe di calore riassuntive ed editor dei report drag-and-drop.

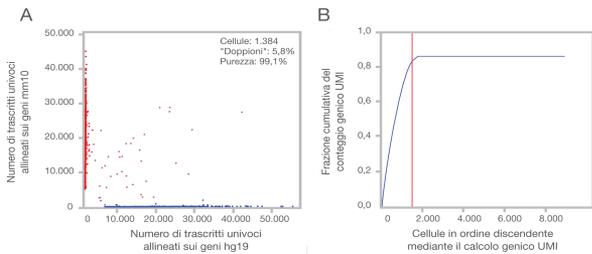
\* Calcoli dei dati in archivio. Illumina, Inc. 2015.

## Elevata qualità dei dati

Per dimostrare l'elevata qualità dei dati del sequenziamento dell'RNA di singola cellula ottenuta con la soluzione Bio-Rad Single-Cell Sequencing Illumina, gli esperimenti *in vitro* sono stati eseguiti con cellule mescolate da specie diverse. Le cellule embrionali umane del rene 293 (HEK 293) e i fibroblasti embrionali murini NIH 3T3 sono stati miscelati in un rapporto 1:1 (se non diversamente indicato) e caricati in quattro camere dei campioni di una singola cartuccia Bio-Rad ddSEQM. L'isolatore ddSEQ SCI ha incapsulato e marcato 1.384 singole cellule. I trascritti marcati sono stati elaborati per il sequenziamento della singola cellula utilizzando Bio-Rad SureCell WTA 3' Library Prep Kit Illumina per il sistema ddSEQ e sequenziati sul sistema NextSeq 550 Illumina. I risultati del sequenziamento sono stati analizzati utilizzando l'applicazione SureCell RNA Single-Cell Illumina.

## Rilevamento del gene affidabile in una popolazione eterogenea di cellule

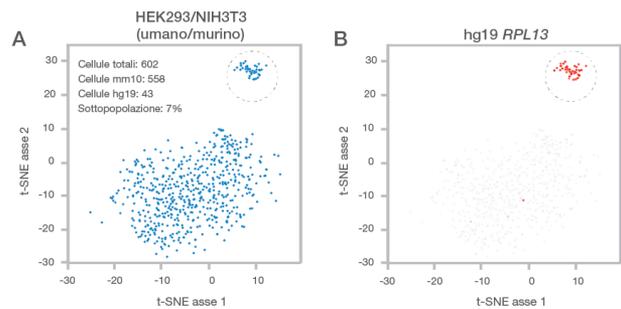
L'analisi dei dati e la rappresentazione grafica dei trascritti univoci umani (hg19) e murini (mm10) rilevati in campioni di specie miste mostrano una bassa percentuale di codici a barre delle cellule (5,8%) contenenti trascritti che possono essere mappati su entrambe le specie e che rappresentano i "doppioni" della cellula (Figura 4A). Questo dimostra l'efficienza dell'isolamento di singola cellula a purezza elevata (99%). Rappresentando graficamente la frazione cumulativa dei trascritti del gene assegnati ai codici a barre delle cellule, il punto di inflessione (linea rossa) viene utilizzato per determinare il numero di cellule marcate rilevate (Figura 4B). Questo indica che un'elevata frazione di trascritti è assegnata a singole cellule.



**Fig. 4. Identificazione affidabile dei trascritti di singole cellule.** **A.** BaseSpace genera grafici relativi al numero di identificatori molecolari univoci (Unique Molecular Identifier, UMI), ossia i trascritti, assegnati ai genomi murini/mm10 (rosso) e umani/hg19 (blu) per ogni codice a barra della cellula. La mappatura univoca dei trascritti sia murini che umani (viola) rappresenta i "doppioni" della cellula. **B.** La frazione cumulativa dei trascritti univoci assegnata ai codici a barre (scala lineare). Il punto di inflessione (linea rossa) determina il numero di cellule marcate rilevate nella corsa di sequenziamento.

## Chiara identificazione delle sottopopolazioni di cellule in campioni misti

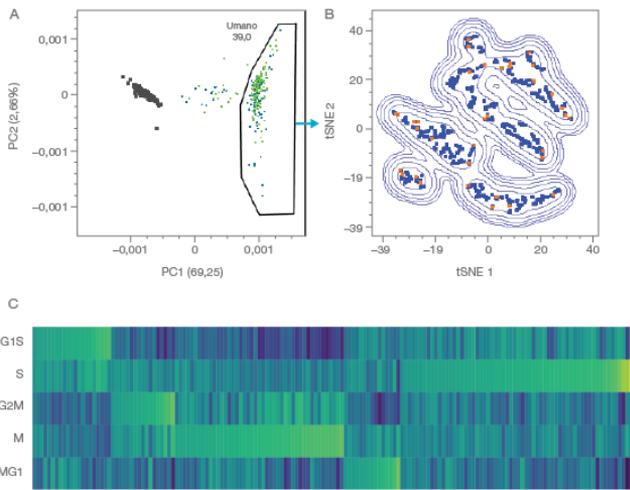
In un esperimento *in vitro* su specie miste, alle cellule murine NIH 3T3 è stato aggiunto un piccolo numero di cellule umane HEK 293. Mediante l'analisi stocastica della distribuzione t integrata con valori simili (t-distributed Stochastics Neighbor Embedding, t-SNE), le cellule umane presenti in questa miscela sono state identificate come cluster distinti e rappresentavano il 7% della popolazione cellulare totale (Figura 5A). Le cellule codificate in base al colore dall'espressione genica umana *RPL13* confermano l'identità della sottopopolazione come cellule umane (Figura 5B).



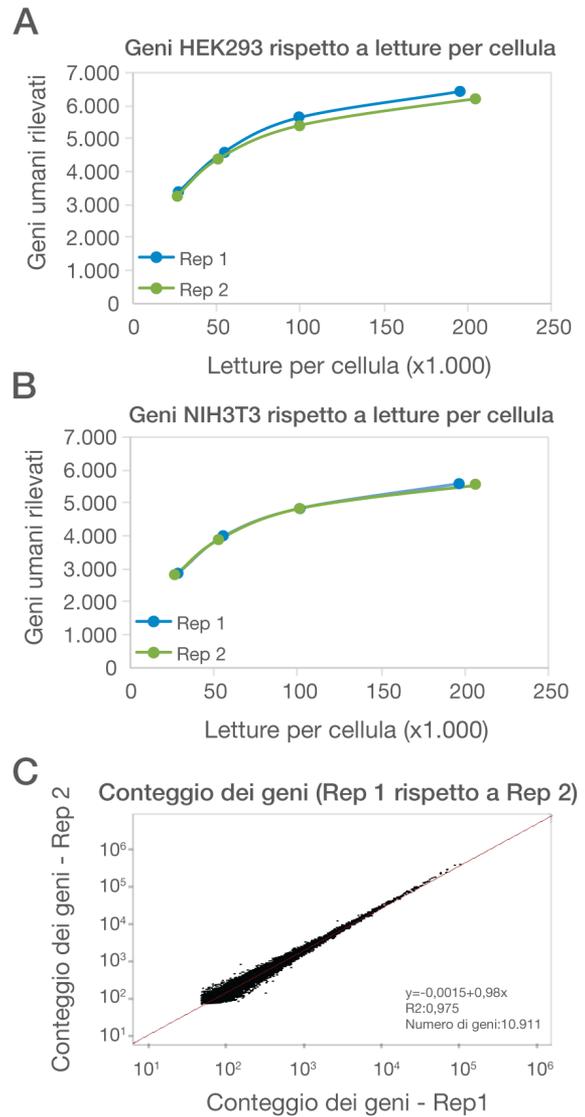
**Fig. 5. Identificazione affidabile della sottopopolazione delle cellule.** **A.** L'analisi mediante l'applicazione SureCell RNA Single-Cell utilizzando l'algoritmo t-SNE di una miscela di cellule NIH 3T3 e HEK 293 identifica una sottopopolazione distinta di cellule. **B.** Le cellule codificate in base al colore dall'espressione genica di hg19 *RPL13* confermano l'identità della sottopopolazione come umana.

## Identificazione dell'eterogeneità delle fasi del ciclo cellulare tramite l'analisi di singole cellule

In un esperimento *in vitro* su specie miste, le cellule HEK 293 sono state mescolate con cellule NIH 3T3 secondo un rapporto 1:1. L'analisi principale dei componenti (principal component analysis, PCA) e il grafico con mappatura a colori dell'espressione genica umana *RPL13* identifica cellule umane (Figura 6A). L'analisi t-SNE del sottoinsieme di cellule HEK 293 rivela cluster distinti di cellule che rappresentano le differenze nei profili dell'espressione genica (Figura 6B). Un'analisi più approfondita dei marcatori associati alle fasi del ciclo cellulare consente un'accurata analisi del ciclo delle singole cellule. Tale analisi si basa sui conteggi dei trascritti univoci dei geni rilevanti, normalizzata per la conta totale di ogni cellula e visualizzata sotto forma di mappa di calore riassuntiva (Figura 6C).



**Fig. 6. Identificazione dell'omogeneità delle cellule HEK293 tramite analisi approfondita della cellula singola con SeqGeq Software.** PCA della miscela di cellule HEK293 e NIH3T3 secondo un rapporto 1:1, mappata in base al colore dall'espressione genica umana RPL13 (A, cellule murine in grigio). Analisi t-SNE del sottoinsieme delle cellule umane in cui cluster distinti rappresentano le differenze nei profili dell'espressione genica (B). La mappa di calore delle singole cellule rivela lo stato del ciclo cellulare in base ai conteggi dei trascritti univoci dei geni (C). L'espressione viene centrata dalla mediana e messa su scala dalla deviazione assoluta della mediana per ciascun ciclo cellulare. L'analisi è stata eseguita usando BaseSpace Sequence Hub e SeqGeq Software.



**Fig. 7. Sensibilità e riproducibilità del rilevamento dei geni su diverse linee cellulari.** **A.** Il numero medio di geni rilevati per cellula è rappresentato graficamente a profondità di sequenziamento da 25.000-200.000 letture per cellula per due replicati delle cellule umane HEK 293. **B.** Il numero medio di geni rilevati per cellula di due replicati di cellule murine NIH 3T3. **C.** Una regressione lineare dell'espressione genica viene rappresentata graficamente per i geni con  $\geq 50$  conteggi, sommati su tutte le cellule HEK 293 in due campioni di replicati.

## Risultati altamente sensibili e riproducibili

I campioni replicati delle cellule HEK 293 sono stati elaborati e sequenziati su un sistema NextSeq 550, come descritto. Le letture di sequenziamento sono state sottocampionate a diverse letture per cellule in un intervallo di 25.000-200.000 letture. Il numero mediano delle letture rilevato per cellula a ogni profondità di sequenziamento mostra un rilevamento del gene altamente sensibile e riproducibile su replicati di diverse linee cellulari (Figure 7A e 7B). I conteggi totali del gene per ciascun gene umano sono stati sommati su tutte le cellule HEK 293. Un adattamento di regressione lineare dei conteggi del gene sommati tra due replicati elaborati su una singola cartuccia ddSEQ M dimostrano ulteriormente l'elevata riproducibilità dei risultati (Figura 7C).

## Riepilogo

Per favorire la comprensione dell'eterogeneità trascrizionale che guida i complessi sistemi biologici, i ricercatori hanno bisogno di metodi scalabili, a elevata processività e di facile utilizzo per generare migliaia di librerie NGS di singole cellule. La soluzione Bio-Rad Single-Cell Sequencing Illumina rappresenta un flusso di lavoro completo sviluppato grazie a una collaborazione di esperti nel settore dell'isolamento delle cellule in base a goccioline e le tecnologie NGS. Questa soluzione permette di svelare nuovi tipi di informazioni su singole cellule semplificando l'analisi di più campioni in parallelo, in diverse condizioni di trattamento e a diversi punti temporali. Le opzioni di analisi dei dati semplici ma molto efficaci mediante l'applicazione SureCell RNA Single-Cell, abbinata agli strumenti per l'identificazione della popolazione e la riduzione dei dati di SeqGeq Software, possono separare popolazioni cellulari eterogenee e individuare le sottopopolazioni di interesse utilizzando i profili dell'espressione genica e gli strumenti per la visualizzazione dei dati. L'analisi di marker dei cicli cellulari permette l'analisi dei cicli cellulari di singole cellule in tessuti complessi. La soluzione Bio-Rad Single-Cell Sequencing Illumina consente l'interrogazione altamente sensibile e riproducibile di trascrittomi di singole cellule da centinaia a decine di migliaia di singole cellule in un solo esperimento.

## Informazioni per gli ordini

Prodotto	N. di catalogo
SureCell WTA 3' Library Prep Kit (kit da due cartucce)	20014279
SureCell WTA 3' Library Prep Kit (kit da sei cartucce)	20014280

## Maggiori informazioni

Per maggiori informazioni sulla soluzione Bio-Rad Single-Cell Sequencing Illumina, visitate la pagina Web:

[www.illumina.com/surecell](http://www.illumina.com/surecell)

[www.bio-rad.com/ddSEQ](http://www.bio-rad.com/ddSEQ)

[www.flowjo.com/seqgeq](http://www.flowjo.com/seqgeq)

## Bibliografia

1. Kolodziejczyk AA, Kim JK, Svensson V, Marioni JC, Teichmann SA. The technology and biology of single-cell RNA sequencing. *Mol Cell*. 2015;58(4):610–620.
2. Shapiro E, Biezuner T, Linnarsson S. Single-cell sequencing-based technologies will revolutionize whole-organism science. *Nat Rev Genet*. 2013;14(9):618–630.
3. Hindson BJ, Ness, KD, Masquelier DA, et al. High-throughput droplet digital PCR system for absolute quantitation of DNA copy number. *Anal Chem*. 2011;83(22):8604–8610.

Solo a uso di ricerca. Non usare in procedimenti diagnostici.