

Trousse de préparation de librairies SureCell^{MC} WTA 3' d'Illumina^{MD} et de Bio-Rad^{MD} pour le système ddSEQ^{MC}

Jetez une lumière nouvelle sur l'expression génique grâce à une plate-forme de séquençage nouvelle génération sensible, ajustable, économique et de grande capacité

Points saillants

- **Caractérisation sensible et objective des signatures transcriptionnelles**
Détection de gènes reproductible et hautement sensible dans les cellules uniques
- **Flux de travail exhaustif de séquençage de l'ARN des cellules uniques**
Flux de travail entièrement pris en charge conçu en collaboration avec des pionniers de la technologie
- **Séquençage nouvelle génération robuste intégré à l'analyse simple de données**
Séquençage éprouvé d'Illumina combiné à un logiciel d'analyse convivial et uniforme

Afin de tenir la promesse relative à la biologie des cellules uniques, la solution Single-Cell Sequencing d'Illumina et de Bio-Rad combine la technologie Droplet Digital^{MC} (3) hautement novatrice de Bio-Rad aux technologies de SNG d'Illumina de préparation de librairies, de séquençage et d'analyse. Cette nouvelle plate-forme propose un flux de travail exhaustif et convivial de séquençage de l'ARN à partir de cellules uniques (figure 1) pour effectuer des essais contrôlés avec de nombreux échantillons et en fonction de conditions de traitement et de moments précis. Conçu et pris en charge de manière collaborative par des chefs de file de la technologie, la solution de séquençage de cellule unique de Bio-Rad et d'Illumina permet d'analyser les transcriptomes de centaines, voire de dizaines de milliers de cellules uniques en un seul essai. Cette robuste méthodologie ajustable de préparation d'échantillons de SNG à partir de cellules uniques permet à plus de chercheurs d'utiliser la précision et la sensibilité du séquençage de l'ARN pour trouver des réponses qui normalement ne sont obtenues qu'en interrogeant chacune des cellules.

Introduction

Les systèmes biologiques complexes relèvent des fonctions coordonnées de chaque cellule. Les technologies traditionnelles qui fournissent des données en masse sur les transcriptomes sont incapables de révéler l'hétérogénéité transcriptionnelle qui sous-tend cette complexité (1). Le séquençage de l'ARN à partir de cellules uniques permet d'analyser en profondeur l'expression génique, ce qui permet de faire la lumière sur la fonction cellulaire, la progression d'une maladie et l'efficacité thérapeutique (2). Cependant, il demeure difficile de produire des milliers de librairies de séquençage nouvelle génération (SNG) à partir de cellules uniques d'une manière abordable, conviviale et à débit élevé (2).

Isolement ajustable et souple des cellules uniques

Le ddSEQ Single-Cell Isolator (SCI) de Bio-Rad encapsule et sépare les cellules uniques en gouttelettes de moins d'un nanolitre sur une cartouche jetable. Chaque cartouche peut contenir de nombreux échantillons, et de nombreuses cartouches peuvent être traitées en parallèle, ce qui permet d'isoler des centaines, voire des dizaines de milliers de cellules par jour. La lyse cellulaire et l'attribution d'un code à barres à une cellule se produisent à l'intérieur des gouttelettes, ce qui permet de suivre chacune des cellules partout dans le flux de travail. Cette méthode fondée sur les gouttelettes fait fi de la taille des cellules mammaliennes et donne des profils objectifs des populations de cellules différentes (figure 2).

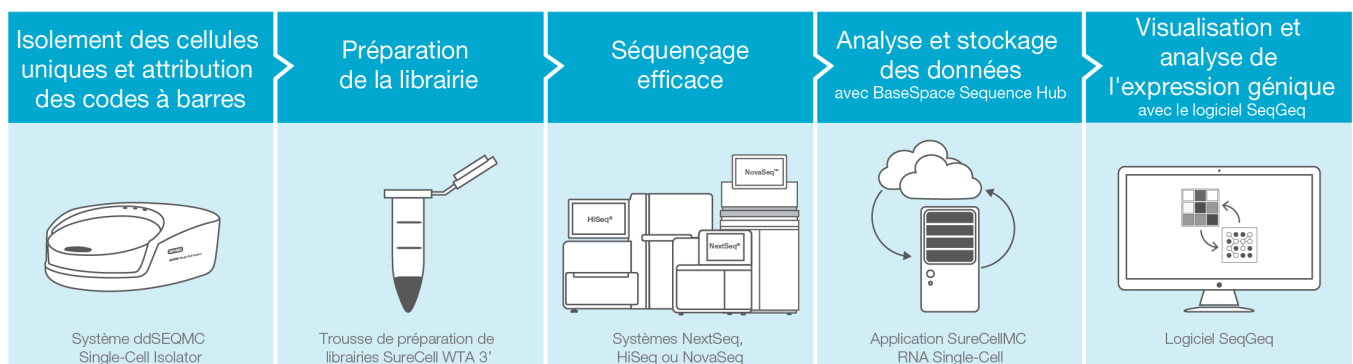


Fig. 1. Flux de travail du séquençage de l'ARN à partir de cellules uniques. Le flux de travail intègre l'isolement cellulaire éprouvé au moyen du système ddSEQ Single-Cell Isolator (SCI) de Bio-Rad suivi de la préparation de librairies à l'aide de la trousse de préparation de librairies SureCell WTA 3' avec la technologie Nextera, le séquençage d'Illumina et l'analyse de données dans BaseSpace Sequence Hub et avec le logiciel SeqGeq Analysis Software de FlowJo, LLC.

Destiné à la recherche uniquement. Ne pas utiliser dans le cadre d'examen diagnostiques.

Le ddSEQ SCI de Bio-Rad peut traiter les échantillons en moins de 5 minutes. L'élimination des longs flux de travail expérimentaux susceptibles de modifier les signatures transcriptionnelles permet de détecter des réponses transcriptionnelles aiguës et d'en faire le suivi dans le temps lors des essais.

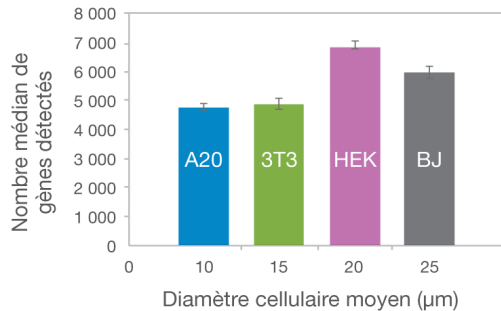


Fig. 2. Séquençage de l'ARN objectif et insensible à la taille de la cellule. Les lignées cellulaires humaines (HEK, BJ) et des souris (A20, NIH 3T3) ont été traitées au moyen de la trousse de préparation de bibliothèques SureCell WTA 3'. Le nombre de gènes détectés constamment élevé montre que la taille des cellules ne limite pas la récolte des transcriptions.

Chimie de SNG avancée pour la préparation de bibliothèques

Une suspension de cellules uniques est mise dans une cartouche ddSEQ M, et le ddSEQ Single-Cell Isolator encapsule les cellules et leur attribue un code à barres. La lyse et l'attribution d'un code à barres se produisent à l'intérieur de chacune des gouttelettes. Les gouttelettes sont brisées, et l'ADNc est regroupé pour la synthèse du deuxième brin. Les bibliothèques sont produites avec un fragmarquage direct de l'ADNc au moyen de la technologie Nextera^{MD}, sans franchir les étapes de séparation ou de préamplification (figure 3). Le fragmarquage est suivi de l'enrichissement de l'extrémité 3' et de l'indexage des échantillons pour préparer jusqu'à 24 bibliothèques indexées et prêtes au séquençage, le tout réalisé avec peu de manipulations.

Cellules uniques isolées, lysées et dotées d'un code à barres

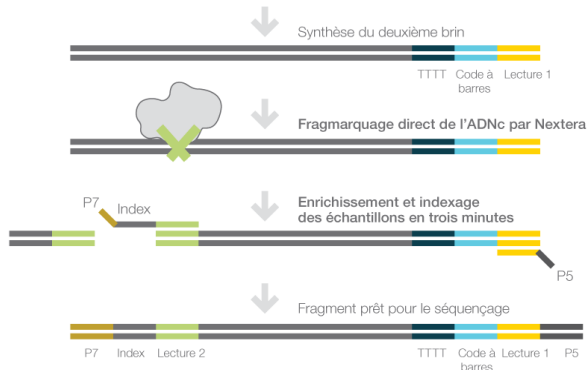


Fig. 3. Aperçu de la trousse de préparation de bibliothèques SureCell WTA 3' pour le système ddSEQ. Une fois que les cellules ont été isolées, lysées et dotées d'un code à barres, l'ADNc est regroupé pour la synthèse du deuxième brin. Les bibliothèques sont produites avec un fragmarquage direct de l'ADNc, suivi de l'enrichissement de l'extrémité 3' et de l'indexage des échantillons.

Séquençage fiable avec les systèmes de SNG d'Illumina

Les bibliothèques de cellules uniques préparées sont chargées directement dans un système MiSeq^{MD}, NextSeq^{MD}, HiSeq^{MD} ou NovaSeq^{MD} d'Illumina pour le séquençage. Ces systèmes de séquençage exploitent la technologie avant-gardiste de la chimie de séquençage par synthèse d'Illumina utilisée dans plus de 90 % de l'ensemble du SNG dans le monde.* Cette technologie répandue permet d'obtenir des données de séquençage cohérentes et fiables dans toutes les plateformes de SNG d'Illumina. Les systèmes de séquençage MiSeq, NextSeq, HiSeq et NovaSeq offrent une vaste gamme d'applications souples et un débit ajustable pour réaliser des études de différentes ampleurs.

Le logiciel de contrôle des instruments facile à utiliser guide les chercheurs dans le chargement des échantillons et la préparation de l'analyse grâce à des interfaces utilisateurs intuitives et il permet de surveiller en temps réel les analyses de séquençage entièrement automatisées sur l'instrument ou à distance.

Analyse des données simplifiée dans BaseSpace^{MD} Sequence Hub

Les données de séquençage à partir de cellules uniques peuvent être instantanément transférées, entreposées et analysées de manière sécuritaire dans le nuage BaseSpace Sequence Hub, l'environnement informatisé de génomique d'Illumina. Le BaseSpace Sequence Hub propose une vaste gamme d'applications. Des outils commerciaux et à code source ouvert prennent en charge une gamme de besoins communs en matière d'analyse de données, comme l'alignement, l'identification des variants, etc. Ces applications affichent des interfaces utilisateurs intuitives à boutons conçues pour être utilisées sans connaissances bio-informatiques particulières.

L'application SureCell RNA Single-Cell a été préparée pour prendre en charge l'analyse de données pour la solution Single-Cell Sequencing Solution d'Illumina et de Bio-Rad. Cette application rend uniforme l'analyse des données allant jusqu'à 96 échantillons répartis dans de nombreuses analyses de séquençage et elle comprend les atouts suivants :

- Paramètres pratiques du contrôle de la qualité du séquençage
- Attribution facile de transcriptions uniques aux cellules uniques
- Rapports et matrices d'expression génique téléchargeables et exportables
- Diverses options d'analyse pour déterminer les sous-populations et les gènes différentiellement exprimés.

* Calculs des données internes. Illumina, Inc. 2015.

Fonctions avancées de visualisation des données du logiciel SeqGeq



Le logiciel SeqGeq est une application de bureau offrant des fonctions avancées d'analyse, d'exploration et de visualisation des données sur l'expression génique des cellules uniques développée par FlowJo, LLC pour la solution Single-Cell Sequencing Solution d'Illumina et de Bio-Rad. Il propose de puissants outils de réduction des données et d'identification de populations. L'intégration directe avec BaseSpace Sequence Hub permet de visualiser et d'analyser des données sur l'expression à l'aide du mappage coloré statistique des cellules uniques, de cartes thermiques sommaires et de fonctions de modification des rapports simplement par glisser-déposer.

Données de haute qualité

Afin de montrer les données de haute qualité du séquençage de l'ARN à partir de cellules uniques obtenues avec la solution Single-Cell Sequencing Solution d'Illumina et de Bio-Rad, des essais *in vitro* ont été réalisés avec un mélange de cellules de différentes espèces. Des cellules rénales embryonnaires humaines (HEK 293) et des fibroblastes embryonnaires de souris (NIH 3T3) ont été mélangés dans un rapport de 1:1 (à moins d'indication contraire) et chargés dans quatre compartiments d'échantillon d'une seule cartouche ddSEQ M de Bio-Rad. Le ddSEQ SC1 a encapsulé 1 384 cellules uniques et il a attribué à chacune d'entre elles un code à barres. Les transcriptions dotées d'un code à barres ont été traitées pour le séquençage à partir de cellules uniques au moyen de la trousse de préparation de bibliothèques SureCell WTA 3' d'Illumina et de Bio-Rad pour le système ddSEQ et elles ont été séquencées à l'aide du système NextSeq 550 d'Illumina. Les résultats de séquençage ont été analysés au moyen de l'application SureCell RNA Single-Cell d'Illumina.

Détection sûre des gènes dans une population hétérogène de cellules

L'analyse des données et le pointage des transcriptions uniques humaines (hg19) et de souris (mm10) détectées dans les échantillons mixtes montrent un faible pourcentage de codes à barres cellulaires (5,8 %) dont les transcriptions renvoient aux deux espèces, lesquelles représentent des doublets cellulaires (figure 4A). Ce résultat indique un isolement des cellules uniques efficace à pureté élevée (99 %). Dans le graphique de la fraction cumulative des transcriptions géniques attribuées aux codes à barres des cellules, le point d'inflexion (ligne rouge) sert à déterminer le nombre de cellules dotées d'un code à barres et détectées (figure 4B). Ces résultats indiquent qu'une fraction élevée de transcriptions sont attribuées à des cellules uniques.

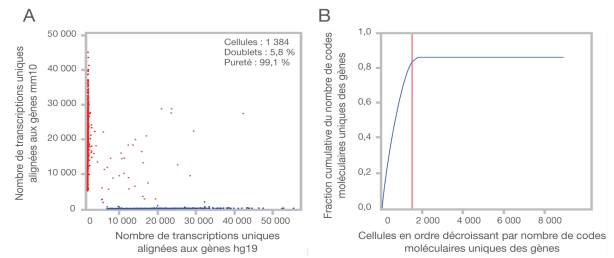


Fig. 4. Détermination sûre des transcriptions des cellules uniques. A.

BaseSpace a généré autant de points que de codes moléculaires uniques (transcriptions) attribués aux génomes humains (hg19, bleu) et de souris (mm10, rouge) pour chaque code à barres de cellule. Les transcriptions uniques qui renvoient à la fois aux humains et aux souris (mauve) représentent des doublets cellulaires. **B.** Fraction cumulative des transcriptions uniques attribuées aux codes à barres cellulaires (échelle linéaire). Le point d'inflexion (ligne rouge) détermine le nombre de cellules dotées d'un code à barres détectées dans l'analyse de séquençage.

Identification évidente des sous-populations de cellules dans les échantillons mixtes

Un essai *in vitro* d'espèces mélangées a été réalisé dans lequel on a injecté un petit nombre de cellules humaines HEK 293 dans les cellules de souris NIH 3T3. L'analyse à l'aide de l'algorithme t-SNE (t-distributed stochastic neighbor embedding) a permis de relever les cellules humaines du mélange comme un amplifiat distinct qui représentait 7 % du total de la population cellulaire (figure 5A). Les cellules colorées en fonction de l'expression génique de la *RPL13* humaine confirment que les cellules de la sous-population appartiennent à l'espèce humaine (figure 5B).

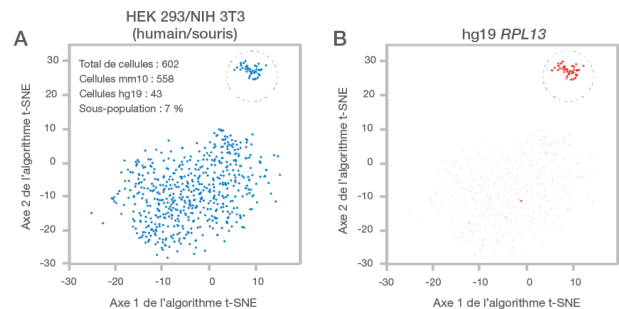


Fig. 5. Identification sûre de sous-populations cellulaires. A. L'analyse du mélange de cellules NIH 3T3 et HEK 293 dans l'application SureCell RNA Single-Cell au moyen de l'algorithme t-SNE permet de relever une population distincte de cellules. **B.** Les cellules colorées en fonction de l'expression génique hg19 RPL13 confirment que les cellules de la sous-population appartiennent à l'espèce humaine.

L'analyse de cellules uniques révèle l'hétérogénéité des phases du cycle cellulaire

Un essai *in vitro* d'espèces mélangées a été réalisé dans lequel on a mélangé des cellules HEK 293 avec des cellules NIH 3T3 dans un rapport de 1:1. L'analyse en composantes principales (ACP) et la représentation graphique par mappage coloré de l'expression génique du gène humain *RPL13* identifient les cellules humaines (figure 6A). L'analyse t-SNE du sous-ensemble de cellules HEK 293 révèle des amplifiats distincts de cellules représentant les différences entre les profils d'expression génique (figure 6B). L'analyse approfondie des marqueurs associés aux phases du cycle cellulaire permet l'analyse précise du cycle cellulaire de cellules uniques. Cette analyse s'appuie sur le nombre de transcriptions uniques des gènes pertinents, normalisé d'après le nombre total pour chacune des cellules et visualisé sous forme de carte thermique sommaire (figure 6C).

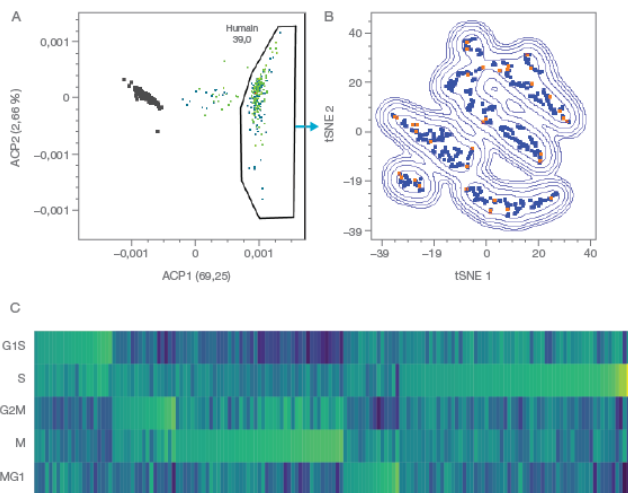


Fig. 6. L'analyse approfondie de cellules uniques par le logiciel SeqGeq révèle l'homogénéité dans les cellules HEK 293. ACP de la mixture 1:1 de cellules HEK 293 et NIH 3T3, représentée graphiquement par mappage coloré de l'expression génique du gène humain *RPL13* (A, cellules de souris en gris). Analyse t-SNE du sous-ensemble de cellules humaines où des amplifiats distincts représentent les différences entre les profils d'expression génique (B). La carte thermique d'une cellule unique révèle le cycle cellulaire en fonction du nombre de transcriptions uniques de gènes (C). L'expression est centrée par la médiane et ajustée en fonction de la déviation de la médiane absolue pour chaque cycle cellulaire. Analyse réalisée à l'aide de BaseSpace Sequence Hub et du logiciel SeqGeq.

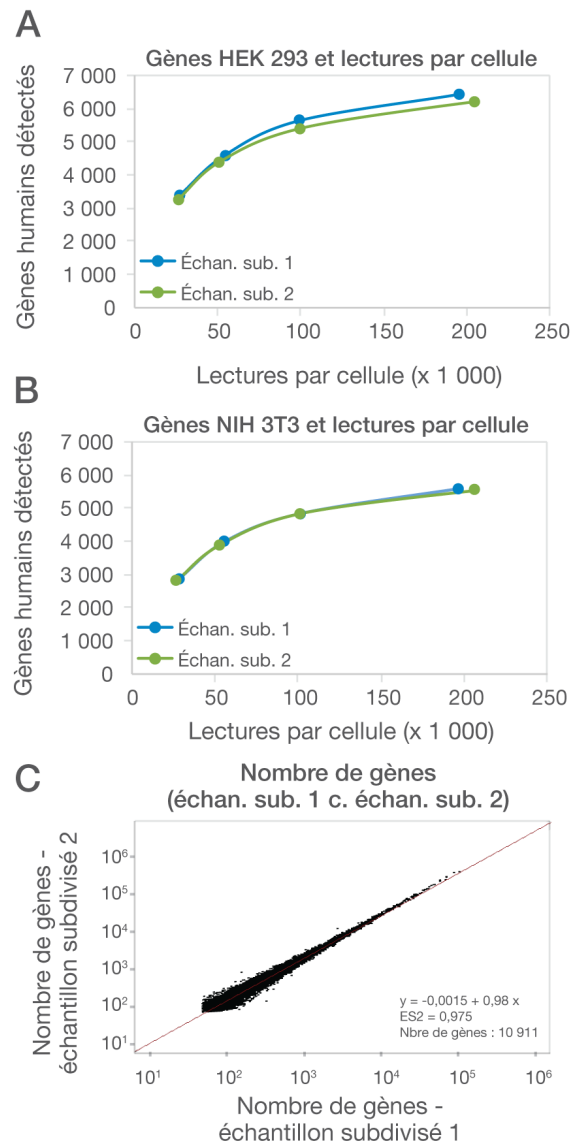


Fig. 7. Sensibilité et reproductibilité de la détection génique dans des lignées cellulaires variées. **A.** Le nombre médian de gènes détectés par cellule est pointé à des profondeurs de séquençage de 25 000 à 200 000 lectures par cellule pour deux réplicats de cellules humaines HEK 293. **B.** Nombre médian de gènes détectés par cellule pour deux réplicats subdivisés de cellules NIH 3T3 de souris. **C.** Régression linéaire de l'expression génique pointée pour les gènes dont le nombre est égal ou supérieur à 50 et calculé pour toutes les cellules HEK 293 dans deux échantillons de réplicats.

Résultats hautement sensibles et reproductibles

Les échantillons subdivisés de cellules HEK 293 ont été traités et séquencés avec un système NextSeq 550, comme décrit. Les lectures de séquences ont été sous-échantillonnées en différentes lectures par cellule, allant de 25 000 à 200 000 lectures. Le nombre médian de lectures détecté par cellule à chaque profondeur de séquençage indique une détection génique hautement sensible et reproductible dans tous les échantillons subdivisés de lignées cellulaires différentes (figures 7A et 7B). Le nombre total de gènes pour chaque gène humain a été calculé dans toutes les cellules HEK 293. Un ajustement de la régression linéaire du nombre de gènes calculé entre deux échantillons subdivisés traités sur une seule cartouche ddSEQ M montre bien des résultats hautement reproductibles (figure 7C).

Résumé

Afin d'approfondir la compréhension de l'hétérogénéité transcriptionnelle sous-jacente aux systèmes biologiques complexes, les chercheurs ont besoin d'une méthode ajustable, conviviale et de grande capacité pour produire des milliers de librairies SNG de cellules uniques. La solution Single-Cell Sequencing Solution d'Illumina et de Bio-Rad est un flux de travail exhaustif issu d'une collaboration des experts de l'industrie en isolement cellulaire à partir de gouttelettes et de technologies SNG. Elle expose de nouveaux types de renseignements sur les cellules uniques puisqu'elle permet d'analyser de nombreux échantillons en parallèle, en fonction de multiples conditions de traitement et à différents moments. Les options à la fois simples et puissantes d'analyse de données de l'application SureCell RNA Single-Cell, combinées avec les outils de réduction des données et d'identification de populations du logiciel SeqGeq, peuvent jeter la lumière sur les populations cellulaires hétérogènes et mettre en évidence des sous-populations d'intérêt grâce aux profils d'expression génique et aux outils de visualisation des données. L'analyse des marqueurs du cycle cellulaire permet d'effectuer une analyse du cycle cellulaire des cellules de tissus complexes. La solution Single-Cell Sequencing Solution d'Illumina et de Bio-Rad permet de réaliser une analyse hautement sensible et reproductible des transcriptomes de cellules uniques à partir de centaines, voire de dizaines de milliers de cellules uniques dans un seul essai.

Renseignements relatifs à la commande

Produit	N° de référence
Trousse de préparation de librairies SureCell WTA 3' (trousse de deux cartouches)	20014279
Trousse de préparation de librairies SureCell WTA 3' (trousse de six cartouches)	20014280

En savoir plus

Pour obtenir de plus amples renseignements sur la solution Single-Cell Sequencing Solution d'Illumina et de Bio-Rad, veuillez visiter le site suivant :

www.illumina.com/surecell

www.bio-rad.com/ddSEQ

www.flowjo.com/seqgeq

Références

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

Destiné à la recherche uniquement. Ne pas utiliser dans le cadre d'exams diagnostiques.