Université de Montréal

Méthodes optiques d'attribution d'identifiants moléculaires à des cellules uniques pour assurer leur traçabilité

par Loïc Binan

Département de pharmacologie et physiologie

Faculté de médecine

Thèse présentée

en vue de l'obtention du grade de *Philosophiae doctor*

Doctorat en génie biomédical

Mai 2019

Université de Montréal
Département de pharmacologie et physiologie, Faculté de médecine

---

Cette thèse intitulée

Méthodes optiques d'attribution d'identifiants moléculaires à des cellules uniques pour assurer leur traçabilité

Présentée par

Loïc Binan

A été évaluée par un jury composé des personnes suivantes

Mathieu Dehaes
Président-rapporteur

Santiago Costantino
Directeur de recherche

Alexis Vallée-Belisle
Membre du jury

Daniel Coté
Examinateur externe

Résumé

Bien qu'apparues récemment, les technologies de séquençage de cellules uniques ont déjà largement illustré l'immense variabilité observable entre les cellules d'un tissu. Dès lors, nombre de techniques permettant de caractériser un échantillon biologique évoluent rapidement vers l'analyse individuelle de chaque cellule constituant l'échantillon. Cette variabilité est en effet porteuse d'informations déterminantes qui sont perdues lorsqu'une étude est réalisée en moyenne sur des millions de cellules. Or, si l'on exclut la microscopie, la majorité des protocoles applicables aux cellules uniques requiert une homogénéisation de l'échantillon en une suspension de cellules. Dès que la structure de l'échantillon est brisée, toute information concernant des éléments distinctifs facilement observables au microscope tels que la position, la forme, les contacts avec l'environnement, la direction et la vitesse de déplacement de la cellule étudiée est perdue. Ces éléments sont pourtant des descripteurs clefs du développement d'un embryon ou d'une pathologie, des réponses immunitaires, du fonctionnement du système nerveux, de la croissance et de la différenciation cellulaire.

L'objectif de ce travail est de développer une approche qui permette d'associer une identité aux cellules d'intérêt afin d'en assurer la traçabilité tout au long des protocoles expérimentaux auxquels elles seront soumises par la suite. Les informations obtenues sur chacune de ces cellules par observation au microscope pourront ainsi être corrélées à celles obtenues lors d'analyses subséquentes.

Dans un premier temps, notre but a été d'attacher un marqueur fluorescent à des cellules arbitrairement choisies une à une dans une image de microscope. Nous avons pour cela développé « Cell labeling via photobleaching » (CLaP). Cette méthode repose sur l'utilisation d'un laser pour photoblanchir un fluorophore, ce qui génère un radical libre et permet la liaison d'une biotine à la membrane plasmique des cellules. Cette procédure est non toxique, n'affecte pas le transcriptome des cellules visées et le marquage peut être détecté plusieurs jours après avoir été placé. Nous faisons la démonstration de principe de l'utilisation de CLaP en conjonction avec la plateforme Fluidigm C1™ pour obtenir la séquence de quelques

cellules choisies individuellement dans un échantillon. Nous étendons ensuite l'utilisation de cet outil à des échantillons tridimensionnels, ainsi qu'à l'ancrage simultané de plusieurs étiquettes de couleur, et à la génération de liaisons entre les cellules et leur substrat afin de permettre leur isolation.

Nous changeons ensuite de paradigme et, plutôt que d'essayer de reconnaître les cellules d'intérêt après avoir séquencé toutes les cellules de l'échantillon, nous cherchons à les isoler. Ces cellules purifiées restent viables et peuvent ensuite être séquencées, réinjectées, cultivées… Pour cela, nous proposons « Single-Cell Magneto-Optical Capture », scMOCA, une adaptation du précédent protocole qui permet de coller des billes magnétiques à la surface de cellules d'intérêt pour les extraire à l'aide d'un simple aimant. Ces manipulations nous ont permis de purifier sans dommages des cellules reconnues très sensibles tels des neurones primaires et des cellules souches embryonnaires. Nous démontrons les capacités de cette procédure en générant des lignées de cellules choisies pour leur capacité exceptionnelle à réparer les dommages induits à l'ADN et ainsi qu'en purifiant des cellules multinucléées jouant un rôle déterminant dans l'apparition des résistances aux médicaments et les récidives de cancer et finalement en extrayant les premières cellules qui se différencient en adipocytes à partir d'une culture de cellules 3T3.

Mots-clés : Cellule unique, marquage, séquençage, photoblanchiment, tri cellulaire, cellule rare, cancer.

Abstract

Even though they only recently appeared, single cell sequencing techniques have already highlighted huge variability among cells. Since then, numerous techniques arose that allow the characterization of each individual cell from a sample. This variability indeed holds crucial information that is lost when studies average observations across millions of cells. Outside of microscopy, most single cell protocols require the creation of a homogenized cell suspension from the sample. Because the spatial structure of the sample is broken, any information easily obtained with a microscope about shape, position, cell-cell contacts, migration direction and speed is lost. These descriptors are nevertheless key to understanding both embryo and pathology development, immune responses, nervous system functioning and cell growth and differentiation.

The objective of this project is to develop a technique that allows giving an identity to cells of interest to trace them across any protocol they might later undergo. This will allow pairing microscopy generated information with that obtained with any downstream analysis.

Our first goal was to tether fluorescent markers to cells that were individually chosen in a microcopy image. We developed Cell labeling via photobleaching (CLAP), in which a fluorophore is bleached using a laser to generate a free radical that allows binding a biotin to plasma membranes. This procedure is non-toxic, leaves transcriptomes untouched, and the tag can be found for several days. We show a proof of principle of the use of this technique with the Fluidigm C1^{TM} platform to sequence individually chosen single cells. We then extended this new tool for 3-dimentional samples, for the simultaneous use of multiple color stamps, and for sorting cells by binding them to their substrate.

We then considered the problem through a different angle: instead of trying to recognize data originating from single cells of interest after sequencing all cells from a sample, one can try to first isolate these few cells prior to sequencing, reinjecting or further culturing them. To this aim, we propose Single-Cell Magneto-Optical Capture (scMOCA), in which the above protocol is adapted to attach magnetic beads on cell surfaces. Viable cells can then be

extracted with high efficiency and purity with a simple magnet; even from populations of very sensitive cells such as primary neurons or embryonic stem cells. Using this procedure, we generated cell lines selected for their outstanding ability to quickly repair induced DNA damage. We then purified multinucleated cells which are involved in the appearance of drug resistance and in cancer relapse and extracted cells that differentiated into adipocytes faster than the rest of the culture.

Keywords : Single cell, labeling, sequencing, photobleaching, cell sorting, rare cells, cancer.

Table des matières

Liste des tableaux

Liste des figures

Liste des sigles

ADN : acide désoxyribonulcléique

ARN : acide ribonucléique

ARPE19 : adult retinal pigment epithelial cells

B4F : biotin-4-fluorecéine

CU : cellule unique

CLaP : Cell Labeling via Photobleaching

Cy5 : cyanine-5

DIBO : dibenzocyclooctyne

DMEM : Milieu essential minimal de Dulbecco

DTT : Dithiothreitol

EDTA : éthylène diamine tétra acétate

FACS : fluorescence activated cell sorting

LAPAP : Laser assisted protein adsorption by photobleaching

MDCK : Madin-Darby canine kidney

IMCD : Mouse inner medullary collecting duct

NA : ouverture numérique (numerical aperture)

PDMS : polydimethylsiloxane

SCS : Single cell sequencing

U2OS : Human bone osteosarcoma epithelial cells

WGA : wheat germ agglutinin

A mes parents et à ma sœur.

Remerciements

Je voudrais commencer par remercier mon directeur de thèse, Santiago Costantino, pour ses conseils, ses enseignements, sa disponibilité.

Je remercie aussi tous les membres du laboratoire pour leur aide et l'ambiance de travail agréable qu'ils ont créée. Merci en particulier à Joannie, pour ces six années passées à servir la science ensemble dans la joie. Merci aussi à Javier pour son aide précieuse.

Je tiens aussi à remercier tous les gens qui ont contribué à ces travaux, les membres d'autres laboratoires du centre de recherche, le personnel du centre Mikhail Sergeev et Martine Dupuy, les stagiaires d'été Jean-Christophe Pelletier de Koninck et Lana Kashino.

Je voudrais aussi remercier mes amis qui m'ont entouré et soutenu pendant ces quelques années à Montréal.

Finalement, merci beaucoup à ma famille.

Chapitre 1 : Introduction

Une grande majorité des techniques d'analyse des échantillons biologiques repose sur l'utilisation de marqueurs. En microscopie, par exemple, des molécules fluorescentes sont fréquemment utilisés pour rendre visibles les structures étudiées. De même, les études protéomiques telles FACS[1], Cytof[2], Abseq[3], CITE-seq[4], sont basées sur l'interaction entre une molécule sonde et un récepteur caractéristique présent sur les cellules d'intérêt. Cette sonde aura au préalable été associée à un agent de contraste, typiquement une molécule fluorescente[5-6], plus récemment un code barre d'ADN[3-4] ou des isotopes métalliques[2, 7-8], afin de permettre sa détection. Cela présuppose d'abord que l'étude porte sur une caractéristique associée à une molécule déjà connue, et ensuite qu'un marqueur ayant une haute affinité chimique pour cette molécule existe et soit connu. Par exemple, aucun identifiant n'est associé à la position de chaque cellule dans l'échantillon, ou au nombre et à la localisation des molécules ou organelles dans la cellule ni même aux contacts entre cellules ou à leur distance par rapport à des sources de signaux moléculaires[9-10]. De même, le mouvement (la migration) est un paramètre facilement observable au microscope, sans marqueur, qui est déterminant dans le développement d'un embryon, d'une tumeur, ainsi que des réponses immunitaires[11]. De manière similaire, la forme des cellules révèle leur croissance, leur division, leur différenciation et leurs pathologies[12] mais aucune molécule ne permet d'identifier sélectivement toutes les cellules allongées ou ramifiées d'un échantillon. Il apparaît ainsi qu'un défaut de marqueurs biochimiques empêche l'étude de certains paramètres caractéristiques d'un échantillon biologique.

D'autre part, la nature chimique des interactions utilisées empêche l'identification spécifique d'une cellule choisie afin de la suivre dans le temps ou au cours d'expériences subséquentes. Pourtant, depuis quelques années, l'étude de cellules individuelles (CI) permet une meilleure compréhension sur les mécanismes qui assurent l'homéostasie des tissus, et comment celle-ci est rompue dans le cas de maladies. En effet, l'incroyable variabilité cellulaire a jusqu'à récemment été dissimulée dans les travaux réalisées sur des millions de cellules, les grandeurs étant mesurées en moyenne sur tout l'échantillon. Au contraire, l'étude de CI

permet de lever le voile sur les interactions entre cellules, d'analyser la progression de maladies provoquées par la présence de cellules rares ainsi que d'étudier les effets du microenvironnement, des niches, sur les phénotypes moléculaires.

Trois situations en particulier bénéficient de l'étude individualisée des cellules. En premier lieu, le séquençage de cellules individuelles a révélé la grande variabilité intercellulaire au sein d'un même échantillon, établissant l'existence de nombreux nouveaux types cellulaires[13-14] et états d'activation impliqués dans la plasticité du système immunitaire[15]. En second lieu, la considération de chaque cellule par rapport à son environnement immédiat s'est révélée riche en information, allant jusqu'à suggérer que certains types cellulaires devraient en fait être divisés en plusieurs sous types de caractéristiques bien différentes[10, 16]. Finalement, les cellules rares nécessitent d'être étudiées individuellement puisque leurs caractéristiques sont effacées par le plus grand nombre de cellules lorsqu'une étude est réalisée en moyenne sur tout un échantillon.

On appelle cellules rares des cellules représentant moins de 0.1% de la population totale de l'échantillon[17]. Il existe de nombreuses situations dans lesquelles des cellules présentes en infimes quantités jouent un rôle déterminant dans le maintien ou la rupture de l'homéostasie. Elles sont parfois même présentes dans un tissu auquel elles n'appartiennent normalement pas. Le développement d'un cancer ainsi que la formation des métastases représentent sans doute les situations les plus évidentes dans lesquelles de rares cellules vont avoir un effet dévastateur sur l'organisme. Les tumeurs résultent en effet de la prolifération anarchique d'une seule cellule[18]. On retrouve d'autre part dans le sang d'un patient atteint d'un cancer de rares cellules tumorales circulantes. Il s'agit de cellules épithéliales qui ont été libérées à partir de la tumeur d'origine[19] avec laquelle elles partagent donc de grandes similitudes. Certaines de ces cellules vont se fixer à nouveau ailleurs dans l'organisme, provoquant des métastases[20-21]. L'extraction de ces cellules et leur analyse est nécessaire pour établir un pronostic et une personnalisation du traitement que le patient doit recevoir. Cependant, leur rareté les rend difficiles à collecter puisqu'on en trouve environ 5 pour 7.5 millilitres de sang[22], soit une pour un milliard de cellules sanguines environ[23]. L'utilisation de

techniques de capture traditionnelles basées sur la micro fluidique pour trier ces cellules peut aboutir à des taux de faux positifs de l'ordre de 81%[23]. De nombreux autres exemples peuvent être cités tels que des populations de cellules souches rares[24-25], des cellules infectées par des virus ou des parasites, des cellules T anti tumeur[26], des cellules inflammatoires[27] ou des cellules fœtales circulant dans le sang maternel[28-29]. L'extraction de ces dernières permet par exemple d'établir facilement un diagnostic non invasif et donc sans danger pour le fœtus [29-30]. La médecine personnalisée[31] et le dépistage de protéines sur des librairies cellulaires[24, 32], qui est très utilisé dans l'industrie et la recherche académique, sont deux exemples de technologies qui font jouer un rôle déterminant à quelques cellules au sein d'une population.

S'il est aujourd'hui admis que des informations capitales sont perdues lorsque des moyennes sont calculées sur des millions de cellules, étudier efficacement un grand nombre de cellules une à une n'en reste pas moins un défi technologique. Un nombre grandissant de méthodes ont pour objectif la caractérisation des protéines que des CI expriment, l'amplification et le séquençage de leur ADN ou de leurs ARN messagers. Lors de ces analyses, l'échantillon est dissocié pour former une suspension de cellules individuelles. Cette préparation de l'échantillon cause la perte de toute information structurelle, en particulier la position, et donc le proche voisinage, de chaque cellule. Une autre difficulté rencontrée après l'identification des cellules d'intérêt est leur isolation. La majorité des approches utilisant la micro fluidique pour isoler des cellules sélectionnées pour la présence, ou l'absence, de multiples récepteurs à leur surface sont optimisées pour atteindre un haut débit afin d'analyser rapidement des millions de cellules[17]. Cependant, ce haut débit est obtenu aux dépens de la sensibilité de ces protocoles, qui sont alors peu efficaces pour détecter des populations de cellules très rares. De plus, même lorsqu'ils atteignent des résultats décents en termes d'efficacité (proportion des cellules ciblées qui sont réellement capturées) et de sélectivité (proportion de cellules capturées qui étaient réellement des cellules d'intérêt), les échantillons qu'ils génèrent sont très dilués à cause du grand volume de fluide utilisé pour créer un flot laminaire[17, 33-34].

Un encodage efficace de ces indications de position, forme, vitesse et direction de déplacement, ne requérant pas une connaissance préalable des caractéristiques biochimiques propres de la cellule, est nécessaire afin de corréler les données générées par microscopie avec celles obtenues par toute autre technologie d'analyse de l'échantillon impliquant une altération de sa structure tridimensionnelle. Il s'agit là de la problématique traitée dans cette thèse.

L'outil de choix pour la mesure des caractéristiques citées précédemment est la microscopie. Les technologies optiques atteignant sans peine des résolutions inférieures au micromètre, elles sont idéales pour caractériser individuellement des cellules. Néanmoins, l'information spatiale visible au microscope est perdue lorsque les échantillons sont homogénéisés en suspension de CI prêtes à être séquencées.

Aussi, l'objectif de cette thèse est de proposer des solutions optiques pour attribuer des identifiants moléculaires à des cellules vivantes. Ces identifiants ne devront pas dépendre d'une connaissance a priori de la surface des cellules visées, et permettre un suivi des cellules vivantes au cours d'analyses subséquentes.

Ce paragraphe clôt le premier des six chapitres de cette thèse. Il est suivi d'une revue de la littérature des technologies disponibles pour marquer des cellules ou des molécules, publiée sous le titre *Exploiting Molecular barcodes in High-Throughput Cellular Assays* dans le journal SLAS Technology. Un troisième chapitre présente une revue des techniques qui permettent de conserver l'information spatiale lors de la caractérisation biochimique d'un échantillon. Vient ensuite une partie décrivant le cœur de la méthode développée pour répondre à la problématique. Cette section est constituée de l'article *Live single cell laser tag* publié en 2016 dans Nature communications. Un cinquième chapitre présente quelques améliorations non publiées qui ont été portées à la technique. S'en suivent deux parties qui détaillent deux technologies issues de ce premier travail, la première publiée en 2019 dans eLife sous le titre *Opto-magnetic capture of individual cells based on visual phenotypes* et la seconde faisant l'objet d'un brevet. Une discussion termine enfin cette thèse.

Chapitre 2 : Revue des approches existantes pour marquer des cellules ou des molécules : article *Exploiting Molecular barcodes in High-Throughput Cellular assays*

L'étude simultanée d'une multitude de paramètres sur un grand nombre de cellules génère un volume de données conséquent. La caractérisation de la présence, l'abondance ou la localisation réalisée simultanément sur plusieurs molécules nécessite de savoir reconnaître à quelle molécule chaque mesure correspond, et, dans certains cas, à quelle cellule elle est associée. L'approche qui est adoptée est d'attribuer une identité, souvent sous la forme d'une étiquette moléculaire, à chacune des molécules et/ou des cellules dont la présence ou l'abondance est mesurée. Cette étiquette assure la traçabilité de la cellule, et de son information génétique, transcriptomique, protéique, tout au long des protocoles qui vont permettre leur étude. Il existe un grand nombre de techniques permettant de générer, puis de lire ces étiquettes. Nous en avons fait une analyse détaillée dans l'article de revue publié dans *SLAS technology* en janvier 2019 sous le titre « Exploiting Molecular barcodes in High-Throughput Cellular Assays ». Ce chapitre est constitué du texte tel qu'il a été publié, les auteurs de l'article sont Loïc Binan, Elliot A. Drobetsky, Santiago Costantino.

a) Abstract

Multiplexing strategies, which greatly increase the number of simultaneously-measured parameters in single experiments, are now being widely implemented by both the pharmaceutical industry and academic researchers. Color has long been used to identify biological signals and, when combined with molecular barcodes, has substantially enhanced the depth of multiplexed sample characterization. Moreover, the recent advent of DNA barcodes has led to an explosion of innovative cell sequencing approaches. Novel barcoding strategies also show great promise for encoding spatial information in transcriptomic studies,

and for precise assessment of molecular abundance. Both color- and DNA-based barcodes can be conveniently analyzed with either a microscope or a cytometer, or via DNA sequencing. Here we review the basic principles of several technologies used to create barcodes and detail the type of samples that can be identified with such tags.

b) Introduction

The widespread implementation of drug screens by both the pharmaceutical industry and academia has triggered the development of barcoding strategies to significantly increase the number of molecules and samples that can be simultaneously characterized. The advent of sophisticated technological hardware for laboratory automation permits highly multiplexed approaches that greatly reduce time and cost. In this context, molecular tags can be used to specifically label, thereby acting as unique identifiers for, a variety of possible entities including individual cells[35], pooled samples[36-37], macromolecules[4], spatial regions[38], and cell lineages[39]. These molecular tags are designed to label specific cells and molecules and possess biochemical properties that facilitate their identification.

The most widespread labelling approaches use either short oligonucleotides[35] or fluorescent labels[40], as these can yield a large number of distinct combinations. Furthermore, identification of such tags is usually performed with standard equipment where sequencing, or spectral detection, are integrated with high throughput assays. For example, short DNA molecules where each base can take four possible values yield enormous numbers of unique permutations. Indeed a 10-base-pair DNA oligo spans 4^{10} (over a million) different combinations. On the other hand, simple color barcodes based on only 5 different fluorescence molecules (e.g. DAPI, FITC, cyanine3, cyanine5, cyanine7 or any dye with similar excitation/emission spectra) in on/off states can generate 2^5 (32) labels. These commonly-used channels can be detected with standard filters available on most fluorescent microscopes, and their number can be further increased with more specialized hardware as mentioned later in this review. In the case of either short oligonucleotides or fluorescent labels, the number of

attributes that can be simultaneously screened increases as a power of the number of channels, thereby generating large numbers of unique barcodes for multiplexing.

For color labels, two additional encoding dimensions can be incorporated to create barcodes. The first relies on different levels of signal intensity[7] to yield higher numbers of combinations. Indeed, while 5 colors used in on/off states generate 32 labels, using a code consisting of three intensities (no signal, low intensity, high intensity) could in principle generate up to 3^5 (243) labels. The second dimension involves positioning colored molecules on a carrier structure[41], so that their order can be measured. For example, the sequence of colors along a carrier RNA molecule can be used just like DNA bases to generate a code[42]. The use of super resolution microscopy allows precise determination of the position of each fluorescent molecule from which such sequences can be inferred. Instead of RNA carriers, hydrogels have also been used to spatially organize colored molecules, for instance within a bead, to create color barcodes[43-44].

For DNA labels, a large number of different strategies has demonstrated the great versatility of this technology. For example, various pipelines developed for single-cell transcriptomics have incorporated different barcoding methods. Currently, the most wide-spread single cell sequencing technology isolates cells in liquid drops, which need to be tagged before being pooled into one sequencing reaction[35]. Barcoding individual cells is achieved via inclusion of distinct short DNA oligonucleotides into all cDNA sequences during library preparation. Such DNA labels are used to assign each read to a cell of origin[35] during analysis.

Based on a similar approach, cellular samples from different origins can also be barcoded, pooled, and sequenced in a single run. Sequenced DNA molecules include both the genetic information and the barcodes that are used to match sequencing information to a sample[36-37]. Considering the important cost of reagents in sequencing technologies, pooling material is crucial towards reducing cost as well as time.

Short DNA molecules are also used to barcode antibodies and proteins, i.e., to combine proteomics and genomics[4]. This powerful approach permits detection of proteins and epitopes alongside transcriptomic data at the single-cell level. Furthermore, barcodes are also used to

tag the position of cells within a sample prior to tissue digestion[45-46]. Thus, transcriptomic data can be matched with spatial tissue organization and cell distribution. Finally, cells can also be barcoded for lineage characterization where a unique identifier is passed to each cell's progeny, allowing one to track differentiation and migration during developmental studies[39].

In this review, we explore how barcodes have recently been exploited in a wide range of applications. We first focus on the use of cellular tags to recognize cells in next generation sequencing pipelines, and then detail how the same techniques are allowing identification of proteins in a sequencing protocol. We also consider how spatial position can be encoded to be paired with a sequencing read of a sample. Finally, we examine how color is being used to barcode various types of probes such as antibodies, proteins, or small ligands used to label cells or DNA fragments.

c) Barcoding for single cell transcriptomics

The use of oligonucleotides as barcodes has been key to the success of next generation sequencing (NGS) techniques[35, 47-48]. Although details vary among sequencing platforms, short DNA identification sequences are incorporated into primers used for library preparation. Most of these, including Nextera™ primers, can be purchased in versions that include short barcodes. Before sequencing, during library preparation, each cDNA molecule is fragmented and extended from both ends with Illumina's adaptor sequences. When desired, each adaptor sequence can include identifiers that generate up to 384 combinations to identify each well in a plate. After library preparation, the 384 encoded libraries are pooled for sequencing, and resulting reads can still be distinguished[49].

More recent single-cell RNA sequencing (scRNAseq) techniques further increase throughput by exploiting microfluidics to encapsulate cells in liquid drops[50]. These drops are generated by water in oil emulsion, where each droplet replaces a well in a plate (Figure 1A). This approach dramatically increases the number of cells that can be simultaneously processed, i.e., up to several thousand. Each captured cell is assigned an identity through a

randomly generated DNA sequence that is immobilized in a gel bead (or on a solid bead) inside a water droplet (Figure 1B). The size and generation rate of liquid drops are tuned to maximize the number of droplets that contain only one cell and one bead. Barcode synthesis is realized by creating DNA molecules on gel beads one base at a time in a controlled fashion[35]. Beads are randomly split into four equal groups, each of which receives one of the four DNA bases. Beads are then pooled and randomly split again into 4 groups for addition of the next base. This process of pooling and splitting is repeated several times, so that each bead carries multiple copies of the same sequence. The huge number of possible combinations obtained with very few bases (16,777,216 for a 12 bp barcode) guaranties a unique code for each bead with very low probability of two beads associated with the same sequence in a sample of a few thousand cells. Since the synthesis of cDNAs from captured single-cells is performed inside the droplet, and all primers in a bead carry the same barcode, cDNA molecules from a single cell share a unique tag.

As an alternative, Ramani et al[51] proposed the use of a similar combinatorial barcoding method on fixed nuclei, without requiring their individual capture. Cells are digested, and nuclei distributed in a 96-well plate with no more than 25 nuclei per well. In each well, DNA within nuclei is tagged by proximity ligation with a first barcode. Nuclei are pooled and split in 96 well plates again, and a second tag is placed at the extremities of the DNA molecules. The grouping of these subsequently added tags creates unique combinations (9,216 combinations in the case of 96 well plates) that can be used to identify individual nuclei.

Single-cell combinatorial indexing RNA sequencing (sci-RNA-seq), a similar method developed by Cao et al. [52], is also based on splitting and pooling fixed cells. Here cells are fixed, permeabilized, and distributed in multi well plates. Each well is then incubated with a specific poly-T primer that includes a handle (i.e., a sequence common to all primers that enables PCR amplification) and a barcode, and mRNA molecules are reverse transcribed. Cells are then pooled and redistributed in multiwell plates where barcoded cDNA molecules are PCR amplified with primers specific for the handle sequence carried by the poly-T primers from the first step. All PCR primers carry their own barcode. Therefore, all cells carry a combination of

two barcodes, one from the primer used in the cDNA synthesis, and one from the primer used for the PCR amplification. Here again, probabilities ensure that they almost all have a different combination of the two barcodes, which allows reliable identification of individual cells.

Most primers used in single-cell NGS studies use barcodes not only to differentiate cells from each other, but also to identify reads originating from single RNA transcripts. Indeed, in the original droplet sequencing (Drop-Seq) paper[35], barcoded primers also contained a random 8 base sequence, termed unique molecular identifier (UMI), which was different for each primer of the same bead among 65,536 possibilities. The presence of UMIs allows filtering of noise amplification artifacts from real cellular expression levels.

The main limitation of single cell sequencing is related to noise, as low expression transcripts are rarely captured which yields highly variable measured signals. The strategies for creating barcodes in this area are relatively well established, and efforts now focus primarily on improving sequencing noise, coverage, and tissue preparation. Another serious limitation originating from the use of beads to associate barcoded molecules to each cell is that, in order to guarantee that single (not doublet) barcoded beads are enclosed with unique cells in droplets, it is necessary to dilute beads, resulting in loss of large numbers of cells[35]. This is not a problem for cell types that are highly represented in the sample; however, losing the majority of cells from a rare population can become a major hurdle. Other techniques used to associate one cell or one nucleus with one barcode are more limited in their throughput as the number of barcoded cells is then limited by the number of wells in a plaque[51-52].

d) Barcoding antibodies for transcriptomics and proteomics

The simultaneous measurement of transcription and translation has represented a technological challenge for decades. Recently, new methods introduced the idea of generating proteins tagged with DNA to convert protein abundance and localisation into data that can be obtained with NGS technologies[3-4]. This novel use of DNA barcodes brings high throughput to proteomic analyses. The capacity to simultaneously read the proteome and the transcriptome

of a cell is of paramount importance; indeed RNA abundance is not always correlated with protein concentration[53] due to variations in post transcriptional processing[54].

The CITE-seq[4] technique achieves simultaneous proteomic and transcriptomic sequencing using DNA-labeled antibodies (Figure 1C) to tag cell surface proteins. Immunolabeled cells are captured for sequencing and the short DNA barcodes ligated to antibodies are detected as cDNAs originating from individual cells. The manner in which these short DNA barcodes are attached to the immunoglobulins varies between protocols. In CITE-seq, biotin and streptavidin are used, whereas in REAP-Seq[55] the barcode is covalently linked to the antibody to reduce steric hindrance. AbSeq[3, 56] relies on a unique molecular identifier attached to barcoded antibodies, allowing measurement of the abundance of individual proteins in cells. All these approaches are being rapidly accepted and used in various studies on cell surface proteins such as immune receptors[18]. Barcoded antibodies have been used to develop a qPCR assay which allows the correlation between numbers of transfected plasmids, transcripts, and barcoded proteins to be evaluated in single cells[57]. In addition to protein detection, barcoded antibodies are also used to quantify epitopes. Lee et al[58] performed western blots of cell lysates and used DNA barcoded antibodies to count single molecules. As all antibodies presented an antigen associated code, precise quantification was possible with specificity comparable to that of ELISA plates.

Genshaft et al[59] developed a similar technique which employed proteins coupled to DNA strands that share a short complementary sequence at their 3' end. When two proteins bind their targets they co-localize sufficiently to allow DNA barcode hybridization. Each probe serves as a primer for extension of the other. This proximity extension assay (PEA, Figure 1D) that requires the tight co-localization of both probes for extension to occur, increases target specificity since remaining non-specific probes do not interact in a way that allows proximity extension. Barcodes are read using the C1 platform from Fluidigm to obtain the full sequence of all tagged antibodies.

Proximity extension assays have also been used to improve signal quality in FISH experiments. Proximity ligation assay for RNA[60] (PLAYR) is based on two barcoded DNA probes

that hybridize *in situ* to improve the strength and specificity of the signal. When two of these probes hybridize in contiguous regions, they capture a third barcoded probe that is then circularized. The fact that precise localized hybridization of two different probes is required to capture the barcoded circle template dramatically increases the technique's specificity. An amplification step is then performed with the circular structure serving as primer. This rolling circle amplification (RCA) of DNA generates a product than contains several repeats of the barcode, thereby generating strong signal amplification. The probed mRNA is converted into a highly repeated barcode compatible with fluorescent and mass cytometry detection. In applying the above technique Frei et al. used DNA barcoded antibodies, and simultaneously detected all barcodes on 14 channels to show strong correlation between RNA and protein localization[60]. Technologies based on antibody recognition require prior knowledge of protein expression profiles in cell samples. Indeed, protein levels are critical to an accurate study based on protein-antibody interaction. More straightforward whole transcriptome sequencing experiments are less sensitive to this as they consider all available genetic information. Furthermore, specificity and affinity of antibodies are highly variable and strongly dependent on experimental conditions, rendering these antibody-based techniques largely experimental at present[3, 61].

A key practical obstacle for single cell sequencing is cost and barcodes have been used to mitigate this by pooling several samples. Barcoded antibodies against ubiquitously expressed proteins with different DNA sequences were employed to tag individual samples[62]. Similarly, Nag et al[63] profiled twenty single nucleotide polymorphisms (SNPs) associated with drug resistance in 463 samples of malaria-infected patients in one sequencing round. This approach reduces the cost by a factor 7, but loses sequencing depth as a trade-off.

An alternative to barcoding antibodies is to use aptamers, which show high specificity for their target molecule. Aptamers consist of RNA, and as such themselves constitute a barcode, i.e., alleviate the need for additional barcoding as in the case of antibodies. Aptamers need only be poly-adenylated to ensure their capture in the next-gen RNAseq workflow[64]. They

are easy to generate using SELEX[65-67] and show binding efficiency and specificity at least equal to that of antibodies[68-69]

The generation of new barcoding sequences cannot be completely random and as such is not necessarily straightforward. While available techniques for generating acceptable sequences are efficient, some constraints must be respected. Amongst these are GC content, homopolymer length, and certain sequences that must be avoided because of their natural presence in a sample or their recognition by a restriction enzyme[70]. These limitations imply that most techniques based on random synthesis of a DNA barcode greatly overestimate the number of useful barcodes that can be generated when groups calculate the theoretical number as an exponential function ($4^{\text{number of bases}}$). Taking into account these considerations, Lyons et al[70] provide a framework for generating billions of acceptable DNA barcodes. Techniques for efficiently tethering a DNA strand to a protein such as an antibody are also being improved[71-72]. Table 1 summarizes key characteristics of each sequencing technique described above.

	Theoretical # barcodes	**Tested # barcodes**	**Processing speed**	**Read depth**	**Doublet rate**	**Capture rate**	**Cost**
Drop-seq[35]	16,777,216	45,000	Thousands per hour	737,000 reads per cell	0.36 to 11.3%	12.8%	7 US¢/cell
sciHiC[51]	# wells in a plate	2,000	Not provided	9,274	4%	100%	Not provided
sciRNAseq[52]	Not provided	15,997	Not provided	32,951	1.7%	100%	20 US¢/cell
Cite-seq[4]	1,024	13	Same as Drop-seq	Same as Drop-seq	Same as Drop-seq	Same as Drop-seq	Not provided
Ab-seq[3]	1,000	2	Same as Drop-seq	Same as Drop-seq	Same as Drop-seq	Same as Drop-seq	Not provided
REAP-seq[55]	65,536	82 antibodies	Same as Drop-seq	20,000	Same as Drop-seq	Same as Drop-seq	Not provided

Tableau I. Comparison of different DNA based barcoding techniques for single cell transcriptomics and proteomics.

e) Barcoding chemical libraries for interaction screening

High throughput screening requires the identification of target-interacting molecules from large candidate libraries. This is rendered difficult by the very limited number of channels offered by fluorescence[73] and mass cytometry[74], even when they are used simultaneously. In theory, DNA barcodes can be employed to easily generate 10^{10} simultaneously usable sequences[75], each of which opens a new experimental channel in which an additional molecule can be observed. This is much more than what can be achieved on FACS platforms[76]. Moreover this high number of barcodes has been exploited to screen major histocompatibility complex multimers[75] and DNA barcoded chemical libraries, where interacting partners can then be identified by a simple PCR [77-79] .

In a similar manner, Pollock et al[80] used phages to carry their barcode. They generated a phage library with each member exposing an antibody fragment (Fab). The DNA sequence encoding the Fab is used as a barcode. They exposed 44 targets to exclusively capture those phages presenting a Fab that had affinity for the displayed targets, whereas other phages were rinsed away. As the captured phage also carried the DNA encoding the Fab, they could identify each interacting Fab by sequencing the phage.

f) Use of barcodes for lineage studies

Barcodes can be used to identify cellular progeny within an organism during normal development, tumor development, or infectious disease propagation[81]. Indeed, including a known short DNA barcode sequence into the genome of a cell of interest ensures that it will be transmitted to progeny, allowing subsequent identification of the latter. As the number of divisions increases, mutations in the genome of the cells appear that create subgroups within the population. It is then possible to establish a genealogy tree of the final population of cells. Bacteria were tagged to study the dynamics of propagation of tuberculosis during the infection of a macaque[82]. The abundance of the subpopulation carrying any given mutation reflects the beneficial effect of said mutation for these bacteria. In yeast, barcodes were used to quantify

the evolution of the relative abundance of 500,000 mutants within a single population[83]. This permitted characterization of evolutionary dynamics after the appearance of beneficial mutations. Moreover; barcode-based lineage studies in bacteria can be exploited to characterize the appearance of drug resistance[84].

Using a library of lentiviruses[39, 85], a number of short DNA sequences can be integrated into cells within an embryo, with different barcodes encoding different cells. Using NGS, these short sequences can be revealed to deduce cell lineage[85]. A modified CRISPR approach based on a homing guide RNA (hgRNA) has also been used to integrate randomly mutating sequences within the genome[86-87]. This method targets nuclease activity to the locus into which the guide RNA is integrated. Therefore, cells can cleave the gRNA locus, which is then repaired in an error-prone manner by non-homologous end joining, generating a new guide RNA, and at the same time mutating the sequence used as barcode. Cell phylogeny can then be inferred from the number and localization of mutations[86, 88]. Interestingly, given the mutation rate, the inclusion of only 6 of these self-mutating barcodes would suffice to uniquely identify all neurons from a mouse. These lineage tracking techniques were coupled with whole cell sequencing workflows to study expression variation during zebrafish development[89]. In this experiment, CRISPR mutations were not random but rather kept under the control of a heat shock-activated Cas9. CRISPR has also been used in Perturb-seq and CROP-seq to introduce changes in selected genes or promoters and characterize their effect on the whole transcriptome. A library of barcoded guide RNAs was used to infect cells. Each guide allowed perturbation of the expression of one gene, which was identified with the barcode carried by the guide RNA[90-92]. Barcodes have also been integrated within viral genomes to track their lineage[93]. Barcoding viruses with 34bp DNA sequences allowed quantification of viral subgroups, and calculation of the reactivation frequencies of the viruses post-treatment.

Finally, color barcodes have also been used for lineage tracing based on Cre-recombinase activity in Brainbow[94]. Cre is able to excise or invert short DNA sequences that are flanked by specifically recognized regions (lox regions). Therefore, infecting cells with distinct fluorescent proteins encoded in a single locus, each flanked by incompatible sets of lox

regions allows the random induction of one of the fluorescent proteins in the cell[94] (Figure 1E). Cre stochastic recombination has been used in very similar ways by various techniques such as BOINC[95] and MultiBow[96].

g) Barcoding spatial information for next generation sequencing

One critical piece of information that can be barcoded, which is otherwise lost in most NGS protocols, is the spatial origin of cells. TIVA[45] allows individual selection of cells within a live microscopy image to be sequenced. To attain such precision, Lovatt et al designed a TIVA tag that enters cells and requires photoactivation to hybridize on polyadenylated mRNA. This tag is biotinylated, which allows downstream extraction of the mRNAs of interest with streptavidin beads. Even though this technique does not reach the read depth of scRNAseq, sequenced cells can be chosen, one at a time, and therefore cellular proximity and contact interactions can be studied[38]. Another method, termed CLaP[46], allows pairing the information generated by single cell sequencing protocols to individual cells in a microscopy image. It uses photobleaching to attach biotin to the membranes of cells that can be chosen based on visible criteria such as shape, migration speed and direction, cell-to-cell contact, or even a characteristic fluorescent signal present within the cell. The biotin can then be targeted with a fluorescent streptavidin. Color tagged cells can be recognized on a Fluidigm C1 sorting chip by epifluorescence imaging, and the whole transcriptome of spatially chosen cells can be evaluated with the typical read depth of NGS techniques.

The other approach for tracking the spatial origin of an mRNA is in situ sequencing, which has the unique capacity to reveal transcript location at the subcellular level. Knowing where transcripts are translated could prove very useful towards understanding functional relationships between genes[38]. Barcodes can be used to mitigate the major drawback of this approach, i.e., the limited number of genes that can be simultaneously observed. Barcoding of “padlock probes” is used to increase the number of sequences that can be simultaneously analyzed[97]. Briefly two 20 base pair DNA probe sequences separated by a 50 bp linker are

hybridized with a cDNA target in situ which, after ligation, creates a circular shaped padlock probe. Ke et al.[98] exploited this approach for in situ target sequencing using a known barcode included in the linker region of the padlock probe. In addition to the signal amplification rolling circle products provide, these are also well adapted to in situ sequencing since they remain bound to the target sequence. Each product can be locally interrogated using sequencing by ligation. In their work, Ke et al[98] encoded probes with 4 bp long barcodes, generating 256 combinations. They used these to locate 31 known transcripts in a breast cancer tissue section. Genes were detected with 98.6% efficiency, with a maximum of 90 reads per cell. This maximal limit is due to the fact that sequencing by ligation is based on imaging, and therefore requires sufficient spacing for the sequenced strands to be discriminated in the image. This is a very powerful method to detect and localize RNAs of known sequence, and Larsson et al[99] used a similar approach to locate DNA molecules.

FISH probes can also be spectrally encoded and then detected by super resolution microscopy[40, 100]. Lubeck et al[40] simultaneously identified up to 32 different barcodes using 3 fluorophores. In this system the code is composed of intensity levels for each of the 3 color channels used to encode the probes. Super resolution microscopy allows sufficient resolution to fluorescently encode, detect, and localize all transcripts associated with a single gene[101].

h) Color barcoding of probes

The number of possible colored probes that can be simultaneously used is restricted, since only a limited number of wavelengths can be detected without spectral crosstalk. To overcome this, several techniques are based on beads that each carry a signal in several color channels. The ratio of intensities in the different detection channels within a bead creates a barcode. Nguyen et al[43] used ratiometric loading of gel beads with 5 lanthanide nanophosphors. These have the advantage of being excited by the same wavelength, do not photobleach, and have narrow emission bands. Different combinations of loading ratios provided 1,101 codes. These beads can be assigned an affinity for a biological receptor by

coating them with a probe to use as an alternative to fluorescent antibodies. In a similar approach, Tang et al[102] stained nematodes with beads loaded with a BODIPY fluorophore flanked by two oxazines. The oxazines can be cleaved by simple light excitation, which shifts the fluorescence of the compound to higher wavelengths. The use of different activation times changes the signal ratios between the three emission wavelengths of the compound, as longer illumination increased the ratio of molecules that had their oxazine cleaved, therefore shifting their fluorescence towards longer wavelength. Different regions of the worm were efficiently encoded by simply varying the activation time along its antero-posterior axis. In a similar approach, Han et al[103] developed microbeads loaded with quantum dots which allow excitation of all channels with a single wavelength. In this protocol the code comprises 10 intensity levels in 6 color channels. The gel beads (approx. 1.2 µm diameter) can be loaded with different numbers of quantum dots and conjugated with DNA capture probes.

Alternatively, DNA has been used as carrier of fluorescent dyes for relative intensity barcoding[42]. Here, the fluorescent molecules are carried by a DNA dendrimer which constitutes a code-carrying microstructure of reduced size thereby improving usability. Two color encoding of DNA probes has also been used to increase the number of targets simultaneously detected by FISH[104].

Another key approach to color coding involves spatially organizing fluorescent molecules on a carrier. This carrier can be a gel bead, within which a barcode can be drawn by photobleaching[105]. Also, a DNA strand can be used as a carrier on which a sequence of color tagged RNA hybridizes, creating a colored sequence. This technique, termed nCounter, was used to count mRNA molecules of over 500 genes and shows high sensitivity without amplification[106]. Each DNA strand is made of a capture sequence specific for the target mRNA, and for a backbone on which colored RNA will hybridize. Using an electric field, all DNA backbones can be aligned in the same direction. Imaging then reveals the color sequence associated with each capture backbone, as well as their number.

On a similar note, DNA origami have been employed[107] to accomplish the same barcoding without requiring application of an electric field, allowing use with live samples. In

this approach, the DNA-PAINT structure is employed to spatially organize colored probes into as many as 216 barcodes. These probes are used to stain live yeast, and super resolution microscopy allows the spatial detection of up to 823,543 codes. In addition to not requiring alignment with an electric field, these probes have the key advantage of being significantly shorter (400-800 nm) than nCounter probes (2µm). Another approach uses structured metallic particles to create a reflected pattern that can be encoded. The advantage here is that all fluorescence channels are left available for more classical stainings[108].

In addition to these approaches using ratios and positions to create codes, Hu et al[109] set out to improve the library of available molecules for spectral encoding. They developed a library of polyynes to establish 20 simultaneously detectable light frequencies. These polyynes can be used to tag any protein and detected in 3 states using Raman spectroscopy: absent, low concentration, and high concentration. With this, a theoretical maximal number of barcodes of 59,048 was attained, the highest number for any optical technique to date.

Fixed tissue samples have been stained with up to 66 different antibodies barcoded with DNA and revealed with fluorescence microscopy in a technique termed CO-detection by inDEXing (CODEX)[10]. Each antibody type is associated with a specific DNA oligo which has a common sequence for a complementary primer, a distinct length, and a very particular design. Antibodies are identified in pairs during extension of the complementary strands of their respective DNA barcode using standard fluorescence microscopy. A mix of fluorescently labeled U (green) and C (red) bases is added to the sample to only reveal the two antibodies with sequences having A or G as a first base after the primer during the first imaging cycle. After each image, fluorophores are cleaved, the excess of DNA bases removed, and either A or G is added to the polymerizing strand to select the next unique pair of oligos that will fluoresce in each imaging cycle (Figure 1G). Using this approach, Goltsev et al. managed to perform 36 imaging cycles with good signal to noise ratio to detect 31 proteins.

Finally, the use of antibodies tagged with distinct elemental isotopes (mostly metals) offers a comparable number of tags, as available panels consist of close to 40 markers[8]. In this approach cells are vaporised by a plasma into a cloud of elemental heavy ions, originating from

tagged antibodies. Time of flight measurements allows identification of each element present in the volatized material and their proportion. Mass cytometry can be used either in a configuration where single cells are directed to the plasma one by one, as in a FACS experiment, or with paraffin embedded tissue sections thereby preserving also spatial information[110-111]. Each different element bound to an antibody thus behaves as a barcode, and the total number of possible codes is limited by the availability of pure isotopes that can be attached to these proteins. A technology that allows combination of several isotopes on one antibody to create multiple codes has not been developed to date.

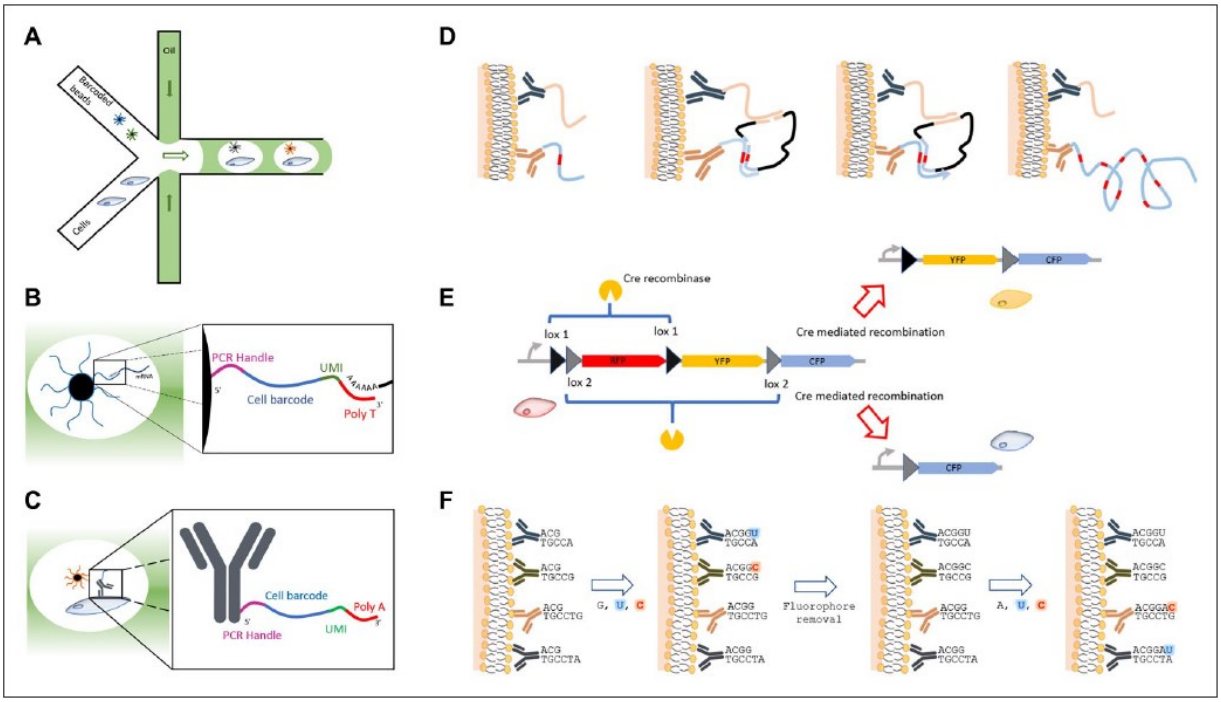


Figure 1. Schematics of the techniques. Single cells are encapsulated with beads and also lysed inside droplets in a microfluidic device[35]. B. Barcoded beads are covered by short DNA oligos containing: a PCR handle to hybridize primers during library preparation, a randomly polymerized sequence of 10-12 bp to barcode each bead, another random 8 bp sequence different in all oligos of the same bead, and finally a poly T sequence of 30 bases to hybridize the poly-A tails of mRNAs. C. Antibodies can be similarly barcoded; here oligo contains a poly-A tail to hybridize as mRNAs originating from the cells do. D During rolling circle amplification, two different antibodies are tagged with different DNA probes and hybridized with two other short ssDNA molecules. Only if antibodies are colocalized, ssDNA strands can be ligated in situ yielding circular DNA. In a final step, this rolling circle is used to amplify a long DNA product that contains several repeats of the antibody-specific sequence[60]. E. In one of Brainbow versions, a Cre recombinase is used to stochastically

excised pieces from a sequence originally inserted within cells' genome and encoding for three fluorescent proteins. By design, only the first (downstream of the promoter) is expressed. Cre specific sequences are flanking these regions in a way that renders them mutually exclusive, resulting in the excision of one, the other, or none of the sequences. After Cre recombinase activation, cells are therefore either still red, or became randomly blue or yellow[94]. F In CODEX, antibodies used to tag cells are barcoded with DNA sequences and their respective primers. A first amplification step with 3 bases, non-fluorescent G, and fluorescent U and C is performed. Sequences containing A are not extended as it is missing from the mix. After image acquisition and fluorophore removal, a second extension is performed with a mix of bases comprising A, and fluorescent U and C. This time, all template sequences containing A or G can be detected in fluorescence, while those containing a C are not extended since G is missing from the mix[10].

i) Conclusions

In this review, we have described many uses for barcodes to identify a variety of objects, from molecules to cells, or even samples. Barcoding offers solutions to many practical problems, including reducing research-associated costs. Moreover, the ability to multiplex allows correlations to be established between biological phenomena in a single run, i.e. obviating the need for separate experiments. From a more academic point of view, although further extensive research is required, barcoding harbours great promise for encoding spatial information and for providing revolutionary methods of precise molecular quantification.

Two main tools are being investigated to barcode information, i.e., the use of synthetic DNA sequences, and fluorescence. Although the former requires sequencing, and therefore sample destruction, DNA tags provide higher numbers of possible combinations, hence more channels that can be simultaneously studied.

Several avenues remain to be explored more deeply. First, even though historically linked with barcode generation, fluorescence suffers from limitations in detection sensitivity. To address this, dyes with sharper excitation/emission need to be synthesized. Also, using colored microstructures to create color sequences that mimic DNA sequences greatly increases the number of barcodes that can be generated with the colors that can already be discriminated with the present technology. A complementary option is to use relative

intensities of different dyes carried by these same microstructures. To further empower these two approaches, barcoded microstructures need to be miniaturized to permit their use in biological samples. Additionally, many groups are focused on creating new techniques to place barcodes on the target cell or structure, such as split and pool encoding, DNA ligation, and antibody or microstructure conjugation, each with its own advantages. More work on these approaches should generate new opportunities.

Finally, despite what may be popularly believed, single cell sequencing techniques only provide means to explore the transcriptome of thousands of single cells, and indeed determining the sequence of a specific single cell chosen in its environment remains a challenge. A minor number of approaches are currently tackling this limitation, and once perfected hold great promise for addressing long standing biological questions where one cell is responsible for great changes such as in organism development, tumor progression, or immunity.

Chapitre 3 : revue des techniques de marquage de la position des cellules

a) Les techniques permettant de corréler information biochimique et origine spatiale d'une cellule

Quelques protocoles replaçant dans leur contexte spatial des données biochimiques générées par les études classiques ont déjà été développés. La méthode la plus ancienne, nommée Fluorescent In Situ Hybridization (FISH), est basée sur l'hybridation in situ de séquences d'ADN qui vont spécifiquement reconnaître et lier les transcrits du gène ciblé. Ces séquences portant une étiquette fluorescente, elles peuvent être identifiées et localisées grâce à un microscope. Cette technique a la particularité de localiser la transcription d'un gène à un niveau subcellulaire. Cependant, elle ne permet pas d'étudier la totalité des gènes exprimés par une cellule, et requiert une connaissance a priori de la séquence du gène à rechercher[112]. L'utilisation des données générées par FISH a cependant permis de créer des cartes 3D des niveaux d'expression de certains gènes dans les organismes les plus étudiés[113]. Ces cartes peuvent être utilisées pour extrapoler la position de chaque cellule dont le transcriptome entier est connu en utilisant les niveaux d'expressions de ces mêmes gènes très courants[114-115]. La résolution de la méthode dépend du nombre de gènes utilisés pour établir la carte, mais elle reste insuffisante pour considérer les contacts entre cellules. De plus, dans le cas de pathologies telles que le cancer et ses métastases, des cellules peuvent migrer à des endroits inattendus. Dans ces situations, leur position ne sera pas extrapolée correctement par cette approche. Alternativement, il a été proposé de séquencer les ARN messagers directement à l'intérieur des cellules, en utilisant l'hybridation in situ et l'observation de sondes d'ADN[116]. L'ARN messager est d'abord rétro transcrit en ADN qui est ensuite ligaturé pour former un cercle. Une ADN polymérase génère ensuite un brin complémentaire en parcourant plusieurs fois ce cercle, formant un produit qui contient plusieurs répétitions de la séquence initiale. Le produit amplifié est alors fixé dans la cellule pour conserver sa position pendant le séquençage.

La cellule est incubée avec un mélange d'amorces qui contiennent toutes une séquence complémentaire d'une portion connue de l'amplicon. Cette séquence, commune à toutes les amorces, est suivie de deux bases aléatoires (4^2=16 combinaisons possibles). Seule l'amorce possédant les deux bases correspondant à celles qui sont effectivement présentes sur l'amplicon va s'hybrider. Ces amorces sont associées à un code couleur, ce qui permet de déterminer par imagerie la séquence complémentaire des deux bases testées. Les amorces sont alors clivées, et un nouveau cycle d'interrogation est réalisé pour déterminer les deux bases suivantes. La répétition de ce cycle permet de séquencer directement à l'intérieur de la cellule. Les principaux défauts de cette approche sont le temps nécessaire (notamment la durée de l'acquisition d'images) et le faible nombre de transcrits détectables dans chaque cellule.

D'autres approches sont basées sur la photoactivation de molécules à l'intérieur d'une cellule à l'aide d'un laser. Ces molécules peuvent servir soit d'étiquette identifiant ladite cellule en la rendant fluorescente, soit d'outil permettant la capture de ses ARNs messagers. Ainsi, une excitation par laser rend fluorescente une version mutée, originalement non fluorescente, de la GFP appelée PA-GFP. Ce procédé a été utilisé pour trier par FACS des cellules individuellement choisies avant de les séquencer[117]. Cette méthode est compatible avec les procédures classiques de séquençage de cellules individuelles et bénéficie de leur haute couverture de séquençage. Sa principale limitation est la difficulté de capturer un petit nombre de cellules individuelles par FACS ; la méthode ne peut dès lors être réellement appliquée à des cellules uniques puisque de nombreuses cellules doivent être photoactivées pour espérer en capturer suffisamment par FACS. La même molécule a aussi été utilisée pour suivre l'évolution dans le tissu, en microscopie 2-photons, de cellules dans lesquelles la PA-GFP avait été photoactivée [118-120]. *Lovat et al.* ont développé une technique appelée « Transcriptome in vivo analysis »[45] pour affiner la résolution spatiale du séquençage à une unique cellule. Une molécule-piège, biotinylée, pénètre dans les cellules et, après photoactivation, libère une séquence poly-U qui capture la queue polyadénylée des ARN messagers. L'utilisation d'une colonne pour lier la biotine permet dès lors d'extraire l'ARN messager de cellules choisies individuellement par laser. Bien que la résolution spatiale de ce protocole permette

effectivement d'identifier les cellules individuellement, le nombre de molécules d'ARN réellement capturées dans chaque cellule est inférieur à celui obtenu par les techniques de séquençage de cellules individuelles classiques.

Une autre méthode consiste à utiliser une plaque sur laquelle sont immobilisées en damier des amorces d'ADN associées à un code barre. Un échantillon peut alors être déposé et lysé sur cette plaque, pour permettre la capture locale des ARN messagers. Lors de la génération du brin d'ADN complémentaire, la position de chaque molécule sur la plaque est encodée par le code barre inclus dans les amorces[121]. Cette approche permet d'évaluer le transcriptome local, mais cependant sans atteindre une résolution de l'ordre de la cellule.

Le recensement de ces méthodes illustre qu'il faut généralement choisir entre un procédé optique, qui conserve l'information spatiale au détriment du nombre de paramètres simultanément mesurables, ou une approche biochimique, qui offre un nombre plus élevé de canaux d'études mais ne conserve pas l'information spatiale. Cependant, de récentes techniques de multiplexage permettent de mesurer un grand nombre (jusqu'à 40) de paramètres par microscopie[10, 16], permettant de remettre ces données dans le contexte spatial avec un niveau subcellulaire. Deux approches, CODEX[10] et 4i[16], utilisent un encodage fluorescent pour interroger itérativement une quarantaine de canaux correspondant chacun à un marqueur. Les images ainsi générées révèlent la localisation des marqueurs à l'intérieur des cellules, et permettent d'évaluer l'influence de son environnement immédiat sur chaque cellule. Elles sont cependant limitées à des études protéomiques, ne révélant pas les séquences ADN ou ARN des cellules observées et sont limitées à une quarantaine de canaux.

b) La fonctionnalisation de substrats assistée par laser

La technologie dont ce travail s'inspire, baptisée « Laser assisted protein adsorption by photobleaching » (LAPAP), permet d'accrocher localement des molécules à un substrat rigide avec une grande précision spatiale. Elle a été développée précédemment dans le laboratoire et repose sur la génération de radicaux libres lors du photoblanchiment de fluorophores[122-123].

Dans la version originale du protocole, des molécules de biotine-4-fluorescéine sont liées sur un substrat de verre. La fluorescéine est un fluorophore dont le principal défaut devient, pour cette utilisation, une grande qualité : il est très facilement photoblanchi. Une illumination avec un laser suffit, même à faible puissance, à provoquer la formation de radicaux libres. La biotine, quant à elle, est une vitamine qui possède une affinité très élevée pour la streptavidine, puisqu'il s'agit du plus fort lien biologique non-covalent connu[124] ($Kd=10^{-14}$). La biotine et la fluorescéine sont de très petites molécules (de masses molaires respectives 332 et 244g/mol quand celle d'un anticorps est de l'ordre de 150 000g/mol) et diffusent donc facilement dans les échantillons biologiques. Leur fusion, la biotine-4-fluorescéine (b4f) est disponible commercialement. Cette paire de molécules permet théoriquement d'accrocher n'importe quel composé sur une surface en quelques étapes. La biotine, liée à la fluorescéine, est d'abord attachée au substrat (verre, plastique…) par une excitation contrôlée au laser qui provoque la conversion locale du fluorophore en un radical libre dans le volume focal du laser. En positionnant précisément ce volume focal à la surface du substrat, la localisation et la densité de la biotine sur le substrat peuvent être contrôlées avec une grande précision. Lors d'une incubation subséquente, la streptavidine va se fixer spécifiquement et irréversiblement sur la biotine, et avec elle toute autre molécule qu'on lui aura préalablement fusionnée, par exemple un autre fluorophore. La précision de la technique dépend de la focalisation du laser et est donc de l'ordre de grandeur de la longueur d'onde soit autour d'un micromètre, ce qui est inférieur à la taille d'une cellule (7-10 µm pour des cellules primaires).

c) Le photoblanchiment des fluorophores

Le phénomène à l'origine des techniques présentées dans cette thèse est le photoblanchiment des fluorophores. Les fluorophores sont des molécules massivement utilisées en biologie afin de générer du contraste dans les images obtenues par microscopie. Lorsqu'il est éclairé, un fluorophore peut atteindre un état excité en absorbant un photon d'énergie fixe, et donc de longueur d'onde déterminée[125]. Pour générer cette lumière de

longueur d'onde contrôlée nécessaire à l'excitation, on utilise généralement soit un laser soit une source de lumière blanche qui est ensuite filtrée. La molécule fluorescente excitée peut revenir à son état énergétique de base en émettant un photon dont l'énergie est inférieure à celle du photon excitant, et dont la longueur d'onde est donc plus élevée. C'est le processus par lequel la majorité des molécules excitées reviennent à leur état fondamental. La détection du photon émis est utilisée pour générer des images en microscopie en fluorescence. Une fois revenue à son état fondamental, une molécule peut à nouveau être excitée en absorbant un autre photon. Cependant, il arrive que deux molécules excitées entrent en collision, ou qu'une molécule excitée entre en collision avec une molécule de dioxygène. On observe alors la formation d'un radical libre, extrêmement instable et réactif[126]. Ces radicaux libres peuvent réagir entre eux, ou avec une molécule d'oxygène, une molécule de fluorophore excitée, ou toute autre molécule donneuse d'électrons. Ce phénomène, appelé le photoblanchiment du fluorophore, est une limitation critique de la microscopie en fluorescence. Tout d'abord parce qu'une fois photoblanchi, le fluorophore n'est plus excitable et ne générera donc plus de signal utile à la création d'une image, ensuite parce que les radicaux libres génèrent des espèces réactives de l'oxygène, extrêmement toxiques pour les cellules. Leur présence lors d'une acquisition d'image sur un échantillon vivant est néfaste puisqu'elle peut provoquer la mort des cellules. Ce sont des molécules qui sont cependant naturellement présentes dans l'organisme et y jouent un rôle important, notamment dans l'immunité ou dans la génération d'énergie cellulaire à l'intérieur de la mitochondrie[127-128]. Leur présence en quantité excessive est souvent corrélée à celle de maladies telles que le cancer, l'arthérosclérose, le diabète, la maladie de Parkinson, les inflammations et infections[129-130]. Les radicaux libres sont capables d'oxyder ou de se fixer sur un grand nombre de molécules. Deux cibles pour lesquelles ils possèdent une grande affinité sont présentes en grande quantité dans la membrane plasmique : les acides gras polyinsaturés et les protéines[127, 131]. En effet, les radicaux libres réagissent avec les doubles liaisons carbones par la réaction

$$X^{\cdot} + CH2 = CHR \rightarrow X - CH2 - C^{\cdot}HR$$

où X· représente le radical libre (par exemple un fluorophore photoblanchi), et R le squelette d'une molécule (par exemple une chaine insaturée d'un acide gras)[131-132]. Lorsqu'elle se produit sur les chaines d'acides gras polyinsaturés, cette réaction peut provoquer une polymérisation et une rigidification de la membrane qui est toxique pour la cellule[129].

Chapitre 4 : fonctionnalisation de membranes cellulaires par photoblanchiment, article *Live single cell laser tag*

Nous avons adapté LAPAP pour permettre la fonctionnalisation de membranes plasmiques plutôt que des substrats inertes. Ce travail a été publié en 2016 dans le journal Nature communication, sous le titre *Live single cell laser tag*.

La technique permet d'attacher une streptavidine colorée à la membrane de cellules choisies individuellement et de suivre l'évolution de ces cellules dans l'échantillon. Nous avons montré que la méthode permet de reconnaître des cellules rares sur la plateforme C1 de séquençage de cellules individuelles. Ce chapitre est constitué du texte tel qu'il a été publié. Les auteurs sont : Loïc Binan, Javier Mazzaferri, Karine Choquet, Louis-Etienne Lorenzo, Yu Chang Wang, Bachir el Affar, Yves De Koninck, Jiannis Ragoussis, Claudia L. Kleinman, Santiago Costantino. Les données supplémentaires publiées avec l'article sont rapportées en annexe.

a) Abstract

The ability to conduct image-based, non-invasive cell tagging, independent of genetic engineering, is key to cell biology applications. Here we introduce Cell Labeling via Photobleaching (CLaP), a method that enables instant, specific tagging of individual cells based on a wide array of criteria such as shape, behavior or positional information. CLaP uses laser illumination to crosslink biotin onto the plasma membrane, coupled with streptavidin conjugates to label individual cells for genomic, cell-tracking, flow cytometry or ultra-microscopy applications. We show that the incorporated mark is stable, non-toxic, retained for several days, and transferred by cell division but not to adjacent cells in culture. To demonstrate the potential of CLaP for genomic applications, we combined CLaP with microfluidics-based single-cell capture followed by transcriptome-wide next-generation

sequencing. Finally, we showed that CLaP can also be exploited for inducing transient cell adhesion to substrates for microengineering cultures with spatially patterned cell types.

b) Introduction

Cellular labels are essential components in the toolbox to build our current understanding of biological function. Yet, a versatile, efficient and non-invasive approach to tag individual cells chosen upon observation is still lacking. The vast majority of methods for generating fluorescently labeled cells rely on biochemical characteristics that are common to an ensemble of cells in a sample, and lack the specificity given by imaging. Widely used methods include transfection of genes encoding fluorescent proteins, membrane-permeable dyes or antibody labeling. These approaches do not allow targeting specific cells among a large population of the same type. Furthermore, their efficiency and specificity are highly dependent on stochastic events and molecular affinity properties, often yielding a sub-optimal fraction of correctly labeled cells. Spatially targeted methods, such as single-cell electroporation[133-134], microinjection[135], laser capture microdissection[135-137], or transfection of photo-switchable proteins that change properties upon illumination[118, 138-139] are often invasive, labor-intensive or lack accuracy, rendering them impractical for a wide range of applications[140-141].

Here we introduce a novel laser-based technique, CLaP, for labeling individual cells in culture. Specific cells can be chosen based on their morphological characteristics, dynamic behavior, localization in the sample at a given time, or any visible feature that distinguishes the cells of interest from an ensemble. CLaP allows combining the accuracy and versatility of image-based selection with the high-throughput of automated cell sorting methods, thus permitting experiments that account for cellular context or temporal dynamics, such as transcriptomic profiling preserving spatial information. The method does not require previous knowledge of cell surface markers, uses off-the-shelf reagents, and may be implemented on a standard confocal microscope without hardware or software modification.

c) Results

i) Cell labeling.

CLaP is related to Laser Assisted Protein Adsorption by Photobleaching (LAPAP) [122-123, 142], a method developed to engineer cell culture substrates by creating protein patterns of optical resolution at a high dynamic range of concentrations. In LAPAP, a laser is used to bind fluorescent biotin conjugates to solid surfaces and hydrogels via free radicals generated by photobleaching. Instead of focusing on inert surfaces, CLaP tethers biotin molecules to the plasma membrane of living cells using a low-intensity laser beam (Fig. 2a). Biotin-4-fluorescein (B4F) is added to the cell culture medium and a laser, tuned near the absorption peak of the dye, is then focused on individual cells of choice, generating reactive oxygen species in close vicinity of the plasma membrane that lead to biotin cross-linking (see Supplementary Note 1). Since the entire process occurs in a small region outside the cell, significant phototoxicity is avoided. The irradiated cells are then revealed by incubating the culture with streptavidin conjugates. By choosing among different types of such streptavidin conjugates, cells can be tagged with fluorescence (Fig. 2b-e), electron-dense molecules (Fig. 2f and Supplementary Note 2) or other labels. The procedure can be repeated sequentially using different color Streptavidin conjugates to obtain distinct color tags within the same sample (Fig. 2e). Tethered biotin spreads along the cell surface via lateral diffusion in the plasma membrane, resulting in a relatively uniform cell staining (Fig. 2d).

Biotin tags can be created with high precision at the single-cell level (Fig. 2). The incorporated mark is well suited for monitoring cell location, movement and progeny, since it displays convenient tracking properties: stable, non-toxic, well retained in cells for at least five days (Fig. 2g), and transfers by cell division (Fig 2g) but not to adjacent cells in a population (Fig. 2c-f, g). Moreover, the label is resistant to routine cell culture procedures. Cells tagged with biotin, resuspended from the substrate with trypsin and reseeded were revealed with fluorescent streptavidin and identified among a large population of unstained cells after 24 hs (Supplementary Note 3). Fluorescence becomes fainter in time as cells divide, and possibly as

proteins are degraded, but subsequent generations of fluorescent daughter cells can be found in the dish (Fig. 2g). Streptavidin tags are mostly restricted to the plasma membrane during the first hours; afterwards, they are gradually internalized, forming bright cytoplasmic puncta (Fig. 2g).

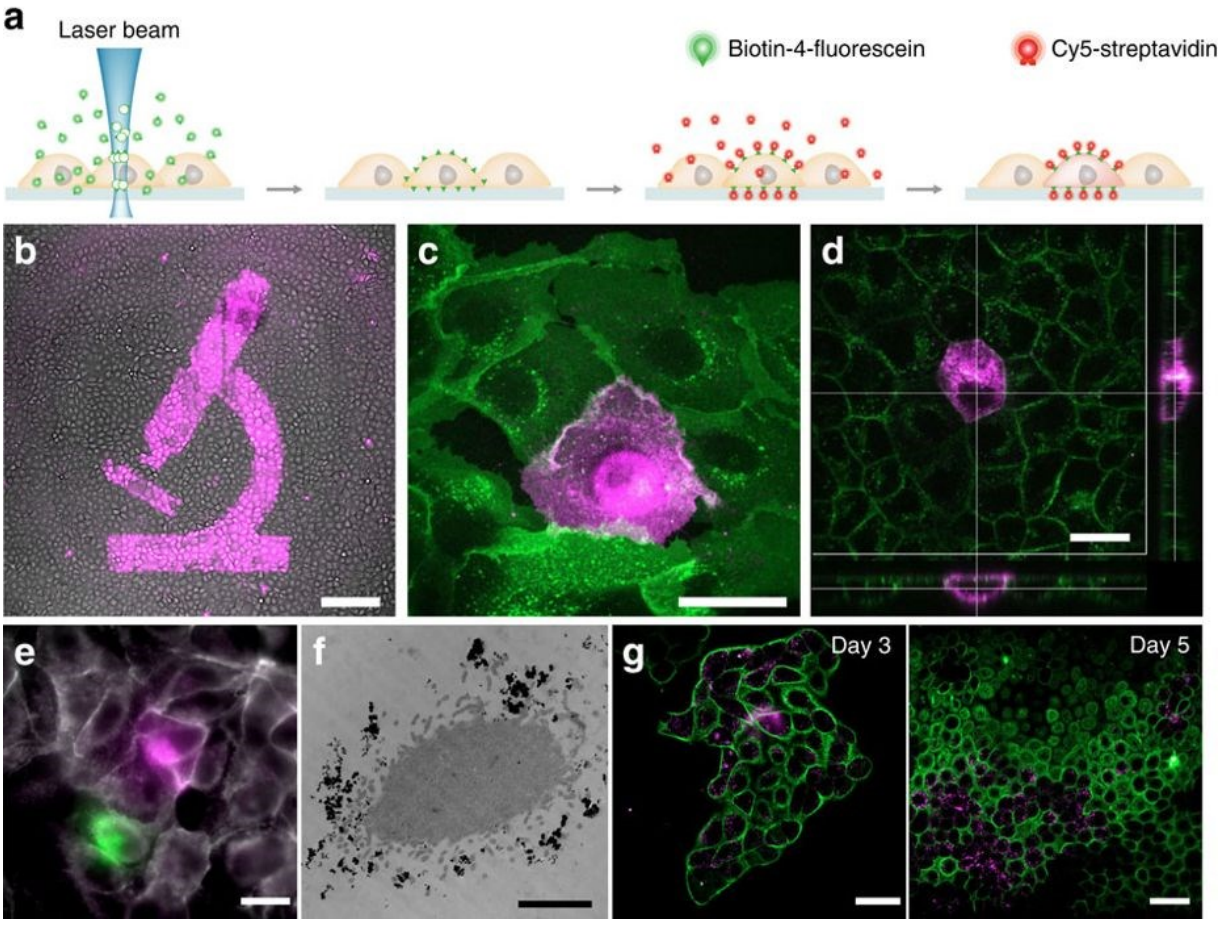


Figure 2. **Cell Labeling. (a)** Outline of the method. Cells are incubated with Biotin-4-Fluorescein, a small molecule that can easily reach the cell membrane, including the space between the glass surface and the cell. A laser beam photobleaches and crosslinks fluorescein conjugated biotin. After rinsing, only illuminated cells retain biotin molecules on their plasma membrane and are revealed with fluorescent streptavidin. Biotin molecules attached to the plasma membrane freely diffuse along the lipid bilayer to yield a rather uniform distribution of fluorophores throughout the cell. **(b)** Examples of labeled cells. Low magnification image of confluent MDCK cells labeled with Alexa-647-Streptavidin overlaid on the bright field image. Scale bar, 200 µm. **(c)** Average confocal projection of a tagged single MDCK cell. The bright circle observed inside the cell boundaries corresponds to Streptavidin bound to the glass, marking the region scanned by the laser. Scale bar, 20 µm. Green corresponds to Wheat Germ Agglutinin-Alexa-488. **(d)** Confocal image and X-Z and Y-

Z projections at day 0 illustrating membrane fluorescence distribution. Scale bar, 20 µm. **(e)** Two-color CLaP obtained by repeating the procedure sequentially and using ARPE-19 cells stained with Alexa647-Streptavidin in purple, Alexa555-Streptavidin in green and WGA-Alexa350 in grey. Scale bar, 10 µm. **(f)** Single labeled cell electron microcopy image, where HRP-Streptavidin was revealed with DAB, mostly concentrated in filopodia. Scale bar, 2 µm. Additional images can be found in Sup. Note 2. **(g)** Fluorescent puncta become visible one day after CLaP and proliferating cells can be tracked for up to at least 5 days. Scale bar, 20 µm.

The amount of biotin-streptavidin complexes bound to the cell membrane scales with the number of molecules that have been photobleached. Hence, the intensity of the cellular stain can be controlled by modulating the total laser energy. We typically deliver 200 µJ over 2 seconds within a small central area of the cells (~100 μm^2, smaller than the cell itself) to obtain a clear staining. Restricting this area to the center of the cells decreases the likelihood of covering neighboring cells, which would result in lower specificity. The average fluorescence intensity of the label is proportional to the laser beam energy, in the range between 150 µW and 400 µW (Supplementary Note 4).

The minimal invasiveness of the procedure is based on photobleaching taking place extracellularly, and cell viability and proliferation do not seem to be affected. To assess cell viability, we used CLaP to tag a square region in two confluent cultures of ARPE-19 cells with streptavidin-Alexa Fluor® 647. Viability staining with Calcein-AM and propidium iodide (Fig. 3a), followed by image segmentation and quantification, showed no increase in cell death after two days, measured at four different time points (Fig. 3b-c). To assess effects on cell proliferation, we tagged isolated cells and quantified the progeny after 3 days (Fig. 3d-e). Once again, no differences were observed between tagged and untagged cells. Finally, the safety of CLaP was confirmed by gene expression profiling of individually isolated cells, where no differences in gene expression was observed upon tagging (see below).

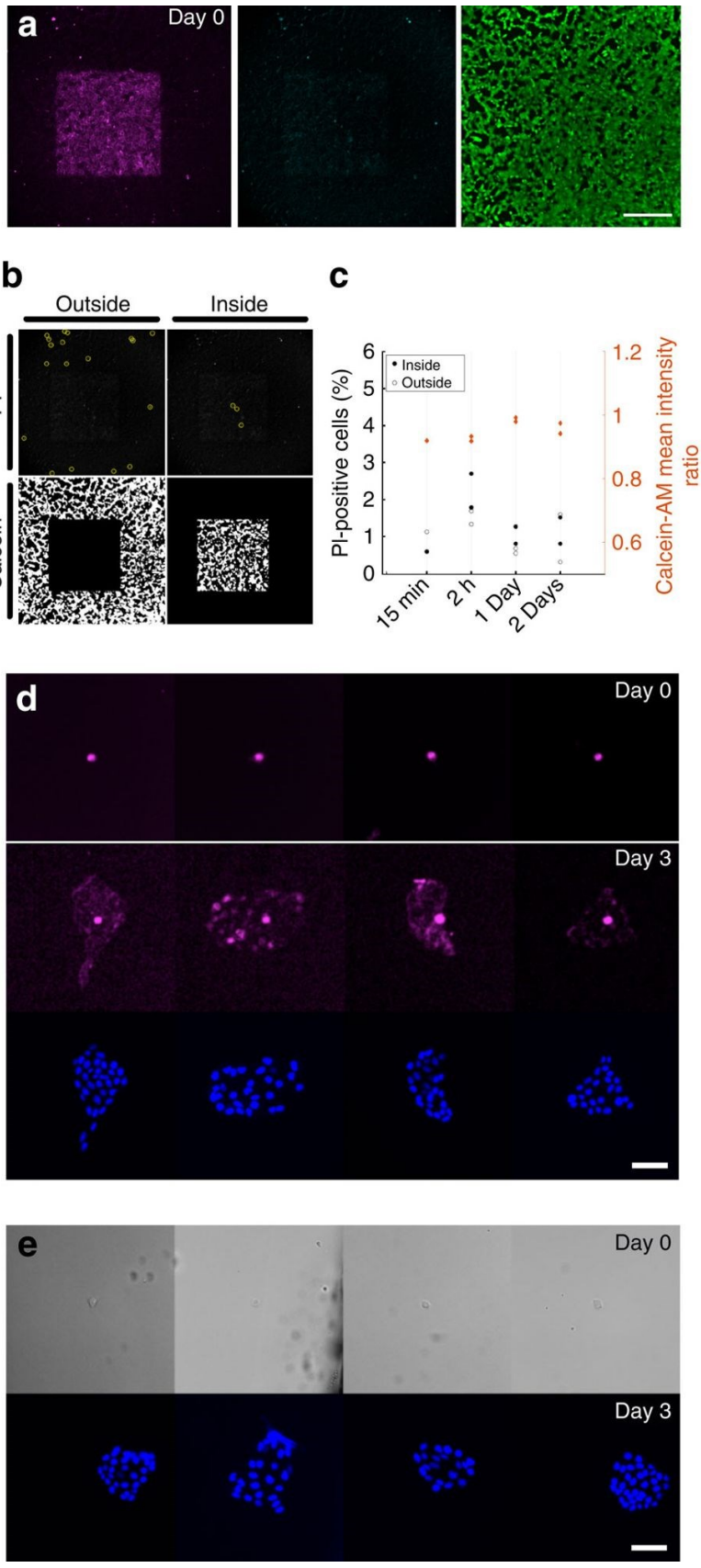

a
Day 0
b
Outside
Inside
PI
Calcein
c
Inside
Outside
PI-positive cells (%)
Calcein-AM mean intensity ratio
15 min
2 h
1 Day
2 Days
d
Day 0
Day 3
e
Day 0
Day 3

Figure 3. Clap labeled cell viability and proliferation. (a) Epifluorescence images of Clap labeled cells. Left (magenta): fluorescent streptavidin channel. Middle (blue): Propidium iodide, indicating dead cells. Right (green): calcein, indicating viable cells. Scale bar 300 µm. **(b)** Images were segmented to assess the fraction of dead cells inside and outside the illuminated region using the PI images. The mean calcein-AM intensity was computed within tagged and non-tagged cells on the segmented images and the ratio between these values was used for quantification. **(c)** Cell viability was quantified at four different time points using both stains, inside and outside the illuminated regions. No significant difference was observed within illuminated ('Inside') and non-illuminated ('Outside') cells. The complete series of images used for this quantification can be found in Supplementary Note 9 **(d)** Isolated MDCK cells were tagged with CLaP using streptavidin-Alexa Fluor 647, and imaged immediately (top panel). After 3 days in culture, cells were fixed, stained with DAPI, and imaged (bottom panel). Single cells proliferated to 28 cells in average with SD = 5.3. Scale bar is 50 µm. **(e)** As a control, non-tagged MDCK cells (top panel) were kept in culture for 3 days, fixed and stained with DAPI. Cells proliferated to 28.5 cells in average with SD = 6.2. Scale bar 50 µm.

ii) Single-cell isolation and genomics.

In order to be studied, tagged cells need to be accurately identified and captured. We first tested CLaP for widely-available fluorescence activated flow cytometry (FACS), a common technology used to sort cells. Three standard gates were defined to count exclusively events originated from isolated viable cells (See Supplementary Note 7 for details). CLaP tagged cells were detected using the Alexa Fluor® 647 signal solely from events going through the three previous gates vs. a dump channel (Fig. 4c).

Next, to demonstrate the potential of the method for single-cell genomic applications, we combined CLaP with microfluidics-based single-cell capture followed by PCR assays and transcriptome-wide next-generation sequencing. We first evaluated the specificity of CLaP and subsequent capture by a co-culture experiment where we tagged only one cell type, captured individual cells, visualized them and analyzed their DNA by PCR. For this, we co-cultured mouse 3T3 fibroblasts expressing mNeonGreen and dog MDCK cells (Supplementary note 5). We exclusively tagged MDCK cells with a far-red fluorescent streptavidin label, targeting approximately 5% of the total number of cells in the dish. Viability test using calcein-

AM/ethydium homodimer before loading the microfluidic device confirmed the non-toxicity of CLaP (Supplementary note 6). We isolated 96 cells using the Fluidigm C1 microfluidic system, among which 5 were CLaP tagged (Fig. 4a). On-chip cell lysis, followed by DNA amplification, allowed us to confirm by PCR the correct species of origin for all tagged cells (Fig. 4b).

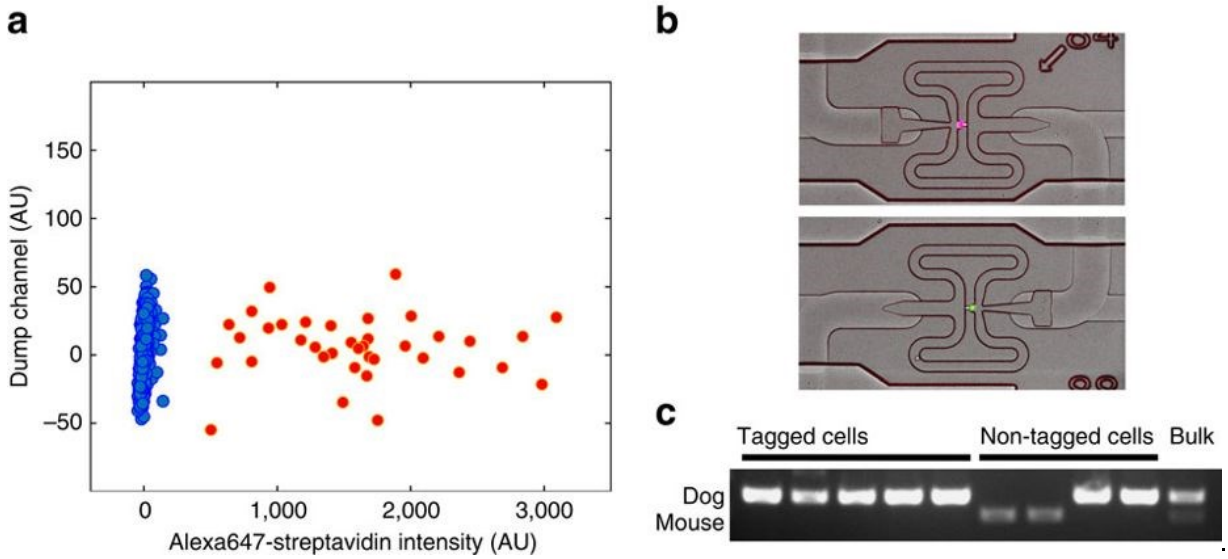


Figure 4. FACS identification of CLaP-labelled cells. (a) After gating on live singlet cells to distinguish laser-tagged cells, the scatter plot shows two populations of cells separated in the fluorescent streptavidin channel, with colour automatically assigned using a hierarchical clustering algorithm (code available as Supplementary Software). Autofluorescence was used to spread the cells in the vertical axis in order to visualize individual cells. (b) MDCK (dog) cells were tagged by CLaP in a co-culture experiment, where ~5% of all cells were targeted. A mix of cells were then individually captured on a Fluidigm C1 chip and visualized. Positive Alexa-647 CLaP tagged cell (magenta) and negative non-tagged cell (green) are shown. (c) PCR amplification confirms species of origin of each tagged cell, isolated with the C1 platform. Non-tagged cells, also isolated with C1, consist, as expected, of a mixture of dog and mouse cells. Rightmost lane corresponds to DNA from a bulk extraction on the rest of the sample. The complete gel and molecular markers are shown in Supplementary Fig. 7.

We then demonstrated that single CLaP positive cells can be isolated from a large population for single-cell transcriptomic analyses, using a human retinal pigment epithelial cell

line. After single-cell capture, RNA-Seq was obtained from CLaP-tagged (n=9) and untagged (n=10) individual cells, in addition to bulk samples consisting of (i) 200 cells pooled, and (ii) cDNA from 5ng total RNA extracted from the bulk culture. We computed a number of quality control metrics to verify that CLaP does not interfere with protocols of sample preparation for transcriptomic experiments. Comparisons of tagged versus untagged cells revealed no significant differences regarding total coverage, GC content and coverage distribution over the genome structure (Supplementary Table 1A and Supplementary Note 8). More importantly, gene expression profiling indicates no major changes associated with the procedure, as unsupervised clustering of samples based on expression profiles (PCA and hierarchical clustering with bootstrapping) consistently groups tagged and untagged cells together (Figure 5 and Supplementary Note 8).

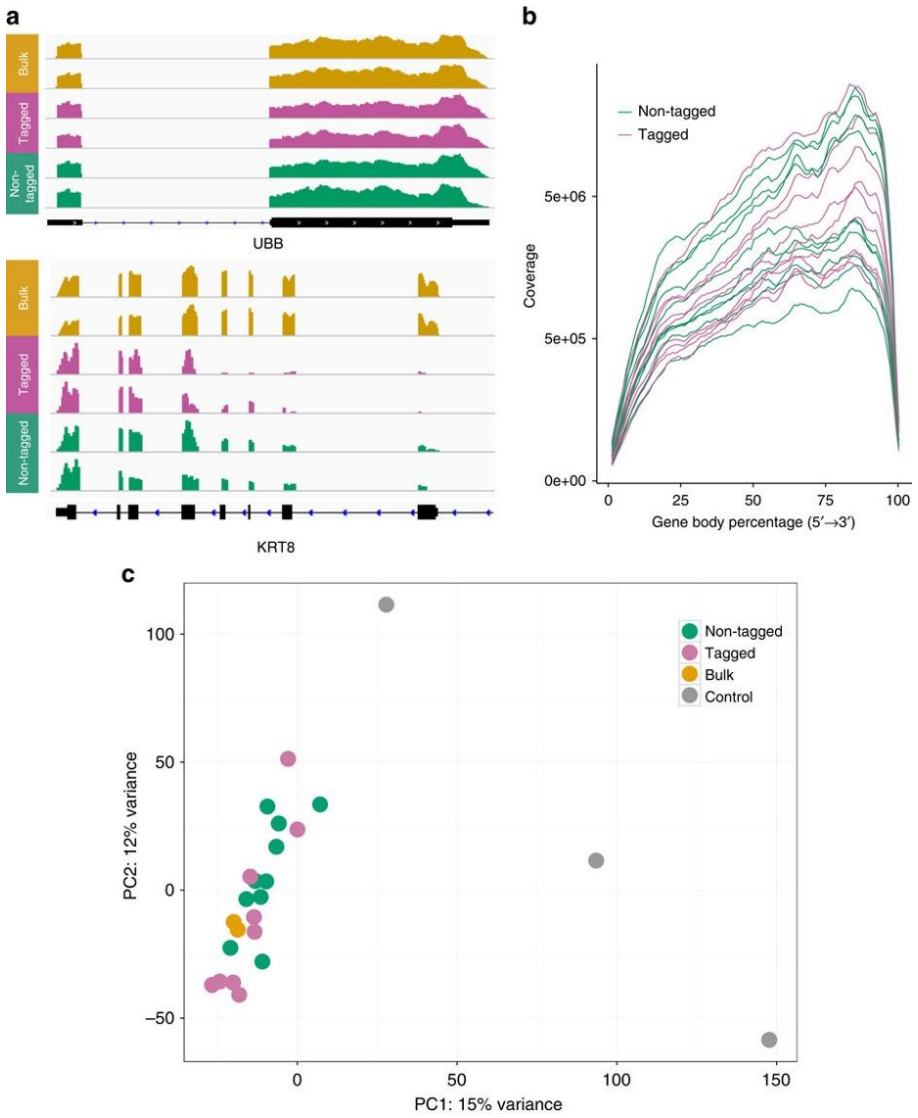


Figure 5. **Single-cell CLaP-labeled RNA transcriptome analysis (a)** Example of RNA-Seq data obtained for one highly expressed gene (UBB) and one RPE marker (KRT8) from bulk (yellow), tagged single-cells (red) and non-tagged single cells (green). **(b)** Coverage uniformity over gene body. Using RSeqC, all transcripts were scaled to 100 nt and the number of reads covering each nucleotide position was computed. The slight 3' bias, reported for SMARTer Ultra Low RNA kit, is expected. **(c)** Global effects of CLaP on cells were evaluated by unsupervised clustering of samples based on expression profiles, which consistently groups together tagged and untagged cells. Different conditions for the clustering, including subsampling genes, bootstrapping or excluding control samples were assessed, with consistent results (See Supplementary Note 8). See Methods for details on control samples.

More subtle effects on gene expression levels were assessed through a differential expression analysis, which showed statistically significant differences in only 24 genes, representing 0.34% of all detected genes (Supplementary Table 1B and 1C). Altogether, these results are consistent with the viability and proliferation assays, indicating that CLaP does not affect regular cell function.

iii) Laser-controlled spatial distribution of cells.

The same photochemistry used to tether biotin molecules to the plasma membrane induces transient adhesions of cells to the substrate, which are resistant to EDTA and Trypsin treatment. As previously shown[143], small fluorescent molecules can diffuse between the glass substrate and the cell membrane, allowing crosslinking of the cells to the cover slip upon photobleaching[143]. These adhesions can thus be exploited to tailor the spatial distribution of cells by automating both laser illumination and sample movement (Fig. 6a-b), and subsequently detaching non-tagged cells using proteases or chelators. Cell adhesions induced by CLaP are transient, and cells keep proliferating and migrating away from the initial region where they were attached, spreading to cover the full field of view of a 10X objective at day 5 (Fig. 6d). The transient nature of cell adhesions and the limited impact of the procedure on cell proliferation are probably due to the choice of CLaP fluorophores and the use of low intensity visible light. The method, thus, constitutes a practical way to select clones of proliferating cells based on visual characteristics, which can be both morphologic and behavioral.

We also tested the possibility of combining different cell types in spatially segregated regions of the substrate, for potential use of the technique in tissue engineering. We first illuminated confluent U2OS cells in a square region, detached non-irradiated cells, seeded ARPE-19 cells in the dish and allowed proliferation. After 24h, we re-illuminated both cell-types to create adjacent squares, and detached the rest (Fig. 6c). Of note, to create cell adhesions that withstand treatment with chelators and proteases, the entire cell surface needs to be illuminated, as opposed to the small fraction required to adsorb tags that diffuse along the membrane.

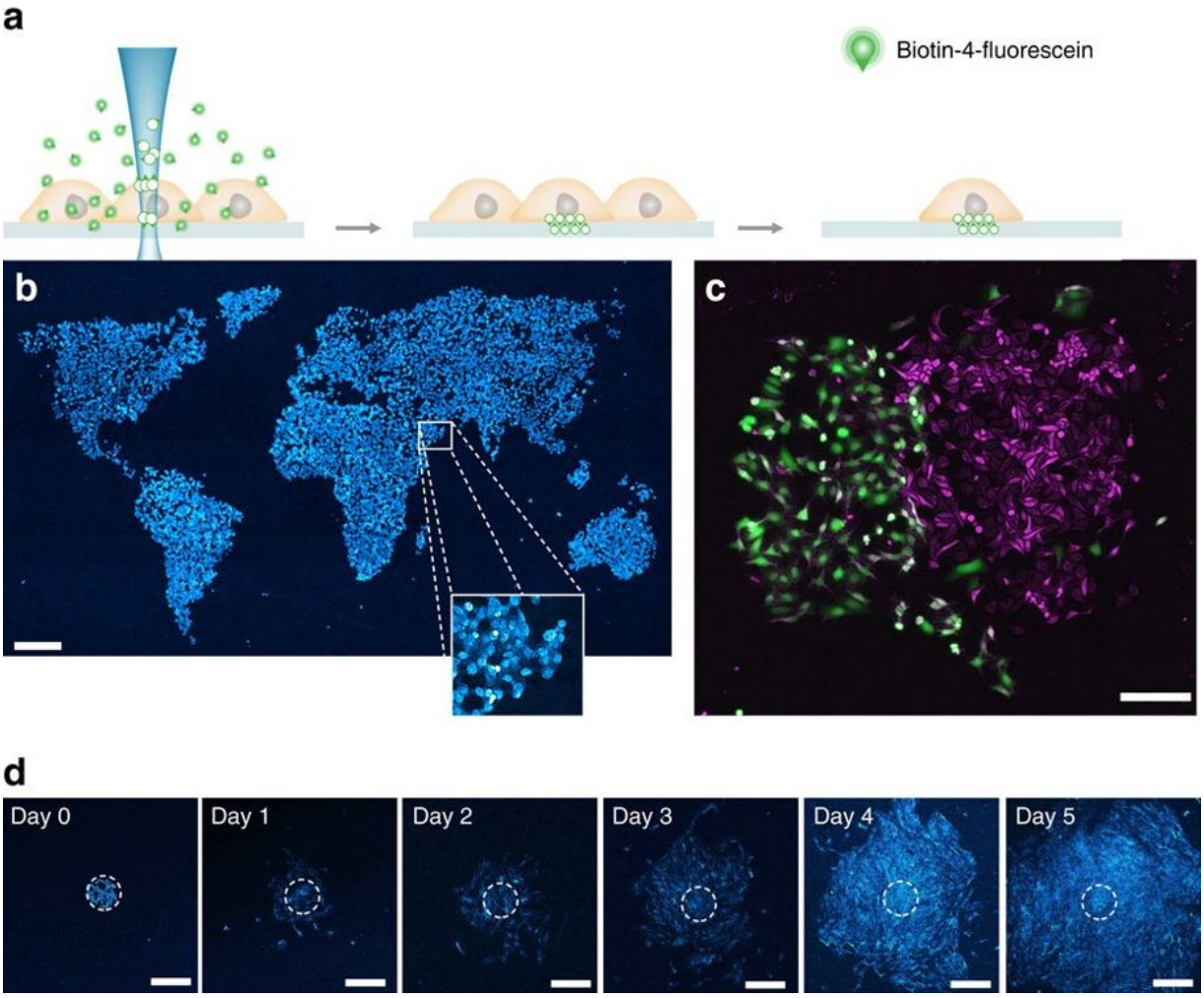


Figure 6. Induced transient cell adhesion. (a) B4F is a small molecule that can easily reach the space between the glass surface and the cell membrane. Reactive species induced by photobleaching B4F create transient adhesions between the cell basal membranes and the substrate. **(b)** Bright field, contrast-enhanced image of a miniature world map created using ARPE -19 cells. Scale bar 400 µm. **(c).** Spatially segregated U2OS cells expressing GFP (left) and bright field contrast enhanced ARPE-19 cells (right). Scale bar 200 µm. **(d)** Cell proliferation after being transiently adhered to the substrate is demonstrated by daily bright field illumination images. Image contrast was enhanced using the method described in the methods. Scale bar is 250 µm.

d) Discussion

In this work, we introduced a method to tag individual cells with a laser, and demonstrated the potential for fluorescence and electron microscopy, as well as FACS and single-cell next-generation sequencing. The most important characteristic of CLaP is the image-based criteria for cell selection, which presents two critical advantages: versatility and potential for automation. This opens the door to experiments that interrogate the transcriptional profiles of cells depending on their microenvironment and spatiotemporal dynamics, permitting to tag, for instance, only fast, large, round, granular, isolated, or distant cells. It also enables to study cell-to-cell communication, in contexts such as immune cell activation or synaptic interactions, where cellular cross-talk induces context-specific molecular changes for which no marker is available. The method allows isolation of cells that have undergone a transient event detectable in a microscopy field. This novel capability is particularly relevant in areas where tissue heterogeneity plays a major role, including development, cancer biology, immune response, stem cells or neurobiology[140, 144].

Recently, a few approaches have been developed for transcriptome sequencing preserving spatial information. *In situ* methods include performing mRNA capture (TIVA[45]) or sequencing reactions (FISSEQ[116]) directly inside intact cells. While CLaP does not have the potential of providing subcellular information or *in situ* expression profiling, it presents several advantages that make it a convenient choice for a number of applications. First, the protocol is short and simple, without requiring special probes, and a standard confocal microscope can be used for tagging, making it a very accessible method. Of course, a dedicated system with more powerful lasers is useful to accelerate the procedure, and programmable motorized stages allow creation of arbitrary illumination patterns. Second, the method is not tied to any particular library preparation protocol or sequencing technology, an important feature in a field that is evolving at a very fast pace. While emerging *in situ* approaches have great potential, practical testing and information on biases and reproducibility are still lacking, and limitations on sequencing depth have to be overcome. CLaP labeled cells, on the other hand, can be analyzed using standard, widely tested capture and sequencing protocols.

Several extensions to the method can be envisaged. Tags are not restricted to fluorescent modalities; here we also used electron dense labels, and bound molecules can easily be extended to magnetic particles, antibodies, nucleotide sequences, drugs or functional macromolecules. Further extensions include color barcoding, to allow separation of more than one group of cells. We have shown that at least two distinct fluorescent tags can be added to a sample (Fig 1e). The maximum number of simultaneous labels is, however, limited by emission cross-talk between image channels, non-specific background of B4F which adds noise to green Streptavidin conjugates, and the diversity of commercially available fluorophores.

We have here demonstrated the use of CLaP on monolayers of cultured cells. The cellular specificity of CLaP is particularly high using cell lines, but it may be slightly decreased in primary cultures, which often include dead cells with permeable membranes and dissection debris that can yield non-specific staining due to adsorbed B4F and Streptavidin. Finally, in tissue CLaP would open a number of exciting possibilities, even if diffusion of reagents through the extracellular matrix for both tagging and rinsing may slow down the procedure. Changing continuous laser illumination for ultrashort laser pulses as a way to spatially confine photobleaching by two-photon absorption in 3-dimensional environments will potentially extend this approach to *ex vivo* and *in vivo* applications.

e) Methods

i) Cell culture.

ARPE-19 cells were cultured in F12 (Thermofisher Scientific, 11765-062) medium supplemented with Gibco Glutamax supplement, 10% FBS and antibiotics. MDCK cells were cultured in DMEM (Thermofisher Scientific, 12561-056) with 10% FBS, Gibco Glutamax supplement and 1% antibiotics and 3T3 and IMCDs in 45% DMEM, 45% F12 with 10% FBS, Gibco Glutamax supplement and 1% antibiotics. Cells were cultured in flasks for amplification and then plated in glass bottom dishes (Mattek p35GC-1.0-14-C) for optimal optical quality.

ii) Single cell labeling.

Cell illumination was performed coupling a 473 nm laser (Laserglow technologies, LRS-0473-GFM-00050-05) to an adapted confocal microscopy module (Thorlabs) mounted on an inverted microscope (Olympus, IX71). In order to determine the localization of the focal spot of the laser, biotin-4-fluorescein (Sigma Aldrich, B9431-5MG) was dissolved at a concentration of 0.04 mg/mL in medium and a drop was dried on a cover slip. Photobleaching of this dried biotin-4-fluorescein upon illumination with the laser allowed precise localization of the focal spot in the microscope field of view.

Dishes containing cells were placed on the microscope and the cells of interest were chosen and illuminated with a power between 100 µW and 350 µW during 2 s using a 0.7 NA objective (Zeiss LD Plan-NEOFLUAR, 441370-9970). This objective ensures enough lateral resolution for aiming the focused laser spot to a single cell. To maximize cell selectivity the illuminated region was restricted to a small central region of the cell membrane, keeping the beam as far as possible from neighboring cells, in order to prevent unspecific tagging. Similarly, the focal plane was set near the top membrane of the cell, which ensures optimal crosslinking efficiency, although this is not a critical issue. After crosslinking, biotin molecules diffuse laterally on the membrane and spread out to the rest of the cell surface. After 3 washes with warm PBS, cells were incubated in streptavidin-compounds at a concentration of 0.05 mg/mL in medium for 15 minutes. The dish was then rinsed 3 times in PBS before imaging.

Depending on the experiment, we have used Alexa-647-Strepatavidin (ThermoFisher Scientific, S-21374); Alexa-555-Strepatavidin (ThermoFisher Scientific, S-32355) or Alexa-488-Strepatavidin (ThermoFisher Scientific, S-11223) all at the same concentration.

iii) Cell viability.

Cells were irradiated by moving the sample at 170 µm/s with a laser power of 240 µW at the sample, using a 0.4 NA objective. Samples were incubated with PI and calcein at four different time points, either immediately after the 15 minutes incubation in streptavidin Alexa 647, 2 hours, 1 day, and 2 days after illumination. (Fig. 3A)

Quantification was programmed in MATLAB by segmenting regions inside and outside the CLaP pattern. First, a binary image was obtained from the Alexa Fluor® 647 image by applying the Otsu algorithm[145], from which only the largest foreground object was kept. The bounding box around it was used as mask to distinguish tagged and non-tagged cells.

For segmenting the image of PI labeled cells, we enhanced objects between 5 and 15 µm wide using a spatial band-pass filter. Local maxima separated at least 15 µm from neighbors were detected, but only those with peak intensities larger than 6 times the standard deviation over the median of all detected peaks were kept (see Fig. 3B). Finally, such peaks were counted inside and outside the CLaP region.

Otsu algorithm[145] was used once more to discriminate between positive and negative calcein pixels (see Fig. 3B).

iv) Cell proliferation.

MDCK cells were plated at very low density, to obtain single cells separated ~1mm from each other. After 2 hours in culture they were tagged with CLaP, incubated with fluorescent streptavidin and imaged immediately afterwards. Cells were kept in culture for 3 days, fixed with paraformaldehyde (PFA) 4% in PBS during 15 minutes, and stained with DAPI. As a control, a second non-labeled dish was kept in culture and cells were monitored daily to check that no nearby cell cluster merged. In order to match the appropriate cluster of cells after 3 days, the original X and Y absolute coordinates of the stained cells at day zero were saved for subsequent imaging of the exact same region.

v) Transient cell adhesions.

Cells were cultured in a 35mm polystyrene dish (Falcon, 35 3001) and incubated with 0.04 mg/mL biotin-4-fluorescein (sigma Aldrich, B9431-5MG) dissolved in medium. The laser was moved at 0.17 mm/s with 240 µW at the sample, and a 0.4 NA objective. Cells were washed quickly with pre-warmed PBS (3 times) and then incubated in 0.05 mg/mL streptavidin-cy5 (Jackson 016-170-084) in medium to simplify the search for the pattern in the dish. After 3 additional washes, cells were placed back in the medium for 4 hours before detachment with EDTA 10 mM (sigma Aldrich E9884-100G).

To obtain spatially segregated cell cultures, a square pattern of U2OS cells was first produced on a 35 mm polystyrene dish (Falcon 35 3001), as described above. ARPE-19 cells were seeded on the same dish at high concentration to reach confluence overnight. Next day, a larger square pattern, including the U2OS cells from the first day, was irradiated to produce a culture of spatially segregated U2OS and ARPE-19 cells.

vi) Cell proliferation after CLaP-induced adherence.

Confluent ARPE-19 cells were irradiated within a 200 µm diameter circular region (Fig. 6, dashed white line) as described above. We incubated cells in EDTA 10mM for 5 minutes, for detachment outside the irradiated region. Attached cells were kept in culture (37 °C, 5% CO_2) during 5 days after CLaP induced adhesion and the sample was imaged daily using a 10XNA0.4 objective with bright field illumination.

vii) Laser and movement automation.

The automation approach used to create spatial irradiation patterns was described in detail elsewhere[122]. A set of instructions describing the stage motor movements and laser power is generated using MATLAB (The MathWorks, Inc.) scripts. These instructions are executed with a program coded in LabVIEW (National Instruments Corporation). Source code is available as supplementary material.

viii) Flow cytometry.

Approximately 10% of cells in a low concentration culture were individually tagged using Alexa-647-Streptavidin. Cells were suspended using Trypsin 0.25% (ThermoFisher, 25200-072), spun and resuspended in PBS containing 5µM Sytox Blue Dead Cell Stain (ThermoFisher, S34857). Flow cytometry was performed with a BD LSR Fortessa cell analyser. We used a SSC-W over FSC-A graph to discriminate doublets, and the Sytox Blue channel to gate on live cells. Propidium iodide fluorescence was used as a dump channel. Colors were automatically assigned using a hierarchical clustering algorithm (MATLAB).

ix) Electron Microscopy.

Cells were cultured on aclar slices to help the preparation of ultrathin slices. Aclar slices were prepared according to previously published protocol[146]. Briefly, 13 mm diameter circles were cut inside an Aclar 22c sheet (Honeywell p5000HS). After several washes and sonication in distilled water, slices were washed in ethanol, then with 10% nitric acid. Aclar slices were then coated with poly D-lysine (100µg/mL, Becton Dickinson, 354210) for 3 hours. Cells were placed in culture on these slices, and staining was conducted in the same condition as previously detailed. Samples were then washed in Sorensen's phosphate buffer and fixed 15 min in 4% paraformaldehyde. After 3 additional washes in Sorensen's buffer, samples were blocked with 1 mL blocking solution (TBS-5% NGS-0.5%) during 45 minutes. Streptavidin sites were revealed by reacting the cells for 15 min at room temperature in 0.3% 3–3´diaminobenzidine tetrahydrochloride (DAB) with ammonium chloride and nickel ammonium-sulfate, and then in the presence of 0.01% H_2O_2. Alternatively, cell were incubated with Streptavidin EM-grade 6nm gold particles, (Electron Microscopy Sciences, 25263) 1:100 in a Lysine-Glycine blocking solution overnight at 4°C. Silver intensification of gold particles was performed for 15 minutes using a Silver enhancement kit (GE Healthcare, RPN491).

Cells were washed in ECS (3 x 5 min), rinsed for 5 min in PB, and incubated in 1% osmium tetroxide (OsO_4) in PB. Finally, the cells were dehydrated through ascending ethanol concentrations and propylene oxide. Sections were then flat embedded in Durcupan ACM. The regions of interest were selected using light microscopy and reembedded in Durcupan ACM

blocks. Ultrathin sections were cut with an ultramicrotome (Leica EM UC7) in 50 nm sections, and collected on Formvar-coated one-slot copper grids and mesh grids. The sections were counterstained with lead citrate before observation on a Tecnai 12 transmission electron microscope (100kV; Philips/FEI) equipped with an AMT V700 camera.

x) Single cell whole genome amplification.

Single-cell whole genome amplified (WGA) DNA was obtained according to Fluidigm protocol "Using C1 to generate libraries for DNA sequencing" (PN 100-7135 H1). Briefly, viable single cells were captured on a 17-25 µm "C1 Single-Cell Auto Prep IFC for DNA Seq" (Fluidigm 100-5764) and visually confirmed using EVOS FL Auto microscope (Life Technologies). The following mixes were made by combining reagents from the Illustra GenomiPhiT V2 DNA Amplification Kit (GE Healthcare Life Sciences, 25-6600-30) and the C1 Single-Cell Auto Prep Reagent Kit for DNA Seq (100-7357); DTT Mix (193.10µL PCR Water, 2.30µL GE Kit Sample Buffer, 2.30µL GE Kit Reaction Buffer, 2.30µL DTT (1M, Fluidigm)), Lysis Mix (19.8µL C1 DNA Seq Lysis Buffer (Fluidigm), 2.2µL DTT (1M, Fluidigm)), Reaction-Enzyme Mix (45µL C1 DNA Seq Reaction Mix, 4.5µL GE Kit Enzyme Mix, 31.5µL DTT Mix). Different mixes were then loaded onto the appropriate IFC inlets.

Inside the IFC, single cells were lysed with 9nL of lysis mix. Total lysis reaction volume was 13.5nL. Next 18nL of C1 DNA Seq Stop Buffer (Fluidigm) were added to the lysis mix. Total stop lysis reaction volume was 31.5nL. Next whole genome amplification (WGA) using the multiple displacement amplification (MDA) method were done at 38°C for 2hrs using 270nL of Reaction-Enzyme Mix. Total MDA reaction volume was 301.5nL. Each single cell reaction was collected into C1 DNA Dilution Reagent (Fluidigm) on a 96 well plate. The final volume per cell was 23uL. In parallel, Tube controls (TC) DNA from ~200 cells (TC_200), a positive control (GE supplied control DNA) and a negative control were similarly processed on a BioRad T100 Thermocycler. Briefly, each TC were lysed with 2uL of lysis mix. Total lysis reaction volume was 3uL. Next the 3uL lysis reaction is neutralized using 4uL of C1 DNA Seq Stop Buffer (Fluidigm). Total stop lysis reaction volume was 7uL. Then 1.05uL of the stop lysis reaction was amplified using 8.95uL of Reaction-Enzyme Mix. Total WGA reaction volume was 10uL. WGA

material was quantified with Qubit dsDNA HS Assay Kit (ThermoFisher, Q32851). The average DNA yield after WGA of single cells, TC200, positive control and negative control was 117n, 437ng, 348ng and 3.35ng respectively.

xi) Identification of species of origin of single-cells by polymerase chain reaction.

Species confirmation was carried out by PCR using Taq (ThermoFisher 18038-042), and 1 µL of the DNA solution prepared with the C1 sorter (Fluidigm) as template, in 25 µL of total reaction volume. The following two primer pairs were used: Cytb1L(5'-CATAGCCACAGCATTCATGG-3'), Cytb1R(5'-GGATCCGGTTTCGTGTAGAA-3'), and Cytb2L(5'-CCTCAAAGCAACGAAGCCTA-3'), Cytb2R(5'-TCTTCGATAATTCCTGAGATTGG-3'), which amplify fragments of 247 nt and 196 nt from the mitochondrial gene Cytb of dog and mouse, respectively.

xii) RNA sequencing.

ARPE-19 cells in culture were isolated as described above. Single-cell mRNA-Seq was done according to the Fluidigm protocol "Using C1 to Generate Single-Cell cDNA Libraries for mRNA Sequencing" (PN 101-7168 H1). Briefly, viable single cells were captured on a 17-25 µm "C1 Single-Cell Auto Prep Integrated Fluid Circuit (IFC)" for mRNA-Seq (Fluidigm 100-5761) and visually confirmed using EVOS FL Auto microscope (Life Technologies). To control for variability in mRNA-seq experiments, 92 ERCC spike-ins (Ambion 4456740) were added to the lysis mix (Fluidigm). The following mixes were made by combining reagents from the SMARTer Ultra Low RNA Kit (Clontech, 634833), C1 Reagent Kit for mRNA Seq (100-6201) and 92 ERCC spike-ins (Ambion 4456740). RNA Spikes mix [0.5uL of 92 ERCC spike-ins (Ambion 4456740), 24.5uL of Loading Reagent (Fluidigm)], Lysis Mix [1µL RNA Spikes mix, 0.5uL RNase Inhibitor (Clontech), 7.0uL of 3' SMART CDS Primer IIA (Clontech), 11.5uL Clontech Dilution Buffer], Reverse Transcription (RT) Mix [1.2uL Loading Reagent (Fluidigm), 11.2uL 5X First-Strand Buffer (Clontech), 1.4uL Dithiothreitol (Clontech), 5.6uL dNTP Mix each dNTP at 10 mM (Clontech), 5.6uL SMARTer IIA Oligonucleotide (Clontech), 1.4uL RNase Inhibitor (Clontech), 5.6uL SMARTScribe Reverse Transcriptase (Clontech)]. These mixes were then loaded onto the

appropriate IFC inlets. Inside the IFC, each single cell was lysed with 9nL of lysis mix. Total lysis reaction volume was 13.5nL. Next mRNA was reverse transcribed at 42°C for 1.5hrs using 18nL of RT mix. Total RT reaction volume was 31.5nL. The full-length cDNA was then amplified through 21 cycles of PCR. Total PCR reaction volume was 301.5nL. Each single cell reaction was collected into dilution buffer (Fluidigm) on a 96 well plate. The final volume per cell was 23uL. In parallel, a set of Tube Controls (TC) using mRNA from ~200 cells (TC_200), 5 ng of purified RNA (Qiagen RNEasy Mini Kit)(TC_RNA) and a negative control (TC_NTC) were similarly processed on a BioRad T100 Thermocycler. Briefly, each single TC were lysed with 2uL of lysis mix. Total lysis reaction volume was 3uL. Next mRNA was reverse transcribed using 4uL of RT mix. Total RT reaction volume is 7uL. Then 1uL RT reaction was amplified through 21 cycles of PCR using 9uL of PCR mix. Total PCR reaction volume was 10uL. The average cDNA yield for single cells, TC200, TC_RNA and TC_NTC were 7.78ng, 134ng, 52.6ng and 29.5ng respectively. On benchtops, cDNA samples were converted to sequence ready libraries using the Nextera XT DNA Sample Preparation (Illumina FC-131-1096) and Index Kit (Illumina FC-131-1002). Briefly, 1.25µL containing 0.375ng of cDNA from every sample were added to 2.5µL Nextera Tagment DNA Buffer, 1.25µL Nextera Amplification Tagment Mix and incubated at 55°C for 10 minutes. Then 1.25µL Nextera NT Buffer was added followed by 5µL of the Nextera Library Amplification Mix, 1.25 µL of Nextera Index 1 primers (N701–N712) and 1.25 µL of Nextera Index 2 primers (S502–S508). The cDNA was then amplified through 12 cycles of PCR and its profile was assessed using the Caliper DNA High Sensitivity LabChip. Equal amount of tagmented cDNA was then pooled from each sample and sequenced on an Illumina HiSeq2500 with paired-end option.

Sequencing runs were processed with Illumina CASAVA software. Reads were trimmed using Trimmomatic v0.32 [147], removing low-quality bases at the ends of reads (phred33<30) and clipping the first three bases in addition to Illumina adaptor sequences using palindrome mode. A sliding window quality trimming was performed, cutting once the average quality of a window of 4 bases fell below 30. Reads shorter than 32 bp after trimming were discarded.

The resulting high-quality RNA-Seq reads were aligned to the human reference genome build hg19 using STAR v2.3.0e[148]. Uniquely mapped reads were quantified using featureCounts v1.4.4 and the Ensembl gene annotation set release 70. Read coverage along gene body was computed using RSeQC[149]. Integrative Genomics Viewer[150] was used for visualization. Multiple quality control metrics (Supplementary Table 1A) were obtained using FASTQC v0.11.2, SAMtools[151], BEDtools[148] and custom scripts. Bigwig tracks for visualization were generated with custom scripts, using BEDtools and UCSC tools.

All the raw data is available on the Sequence Read Archive, accession number SRP069088.

xiii) Analysis of gene expression from RNA-Seq data.

Global expression changes were assessed by unsupervised hierarchical clustering of samples and PCA. To this end, expression levels were estimated using exonic reads mapping uniquely within the maximal genomic locus of each gene and its known isoforms, and normalized (median of ratios), variance-stabilized expression values were derived using DESeq2[152]. Hierarchical clustering was performed using Pearson's correlation as the distance metric and average linkage as the agglomeration method.

To identify specific genes that changed expression upon CLaP labeling, a differential expression analysis between tagged (n=9) and untagged (n=10) cells was performed using DESeq2[152]. Statistical significance was computed using the negative binomial distribution, with the variance and the mean estimated from the data and linked by local regression[152].

xiv) Imaging.

All samples were imaged on an Olympus IX71 microscope (Olympus Corp.) with the appropriate epifluorescence filters and a confocal module (Thorlabs Inc.), or with a FV1000 Olympus microscope. Epifluorescence widefield images were acquired with an Orca Flash 4.0 camera (Hamamatsu Photonics).

xv) Image Processing.

The world map representation image in Fig. 3b is a mosaic of 7x4 10X bright field images. Contrast was enhanced in Fig. 3b and 3c using MATLAB by first calculating the morphological opening of the original image with a 5-pixel circular kernel. The result was subtracted from the original image and a color lookup table was chosen for visualization.

f) Acknowledgements

We thank P. Lemieux for help with electron microscopy, and E. Drobetsky, R. Sieira and S. Chemtob for critical reading of the manuscript. The project was partly supported by the Natural Sciences and Engineering Research Council of Canada, the Canadian Institutes of Health Research, the NSERC CREATE program in NeuroEngineering, and the Canada Foundation for Innovation grants awarded to S.C.

S.C. and C.L.K. receive salary awards from the Fonds de Recherche du Quebec – Santé (FRQS).

L.B. receives a bursary from the Fonds de Recherche en Ophtalmologie de l'Universite de Montreal. K.C. receives a Doctoral Award from the FRQS.

g) Author contributions

L.B., J.M., L-E.L. and Y.C.W. conducted experiments.

L.B., J.M., K.C., S.C. developed software and analyzed data.

L.B., J. M., B.A., Y.D.K., J.R., C.L.K and S.C. conceived experiments.

L.B., C.L.K, S.C wrote the manuscript.

All authors edited and approved the final manuscript.

h) Competing financial interests

The authors declare no competing financial interests.

Chapitre 5 : solutions apportées à certaines limites de la technologie

Nous présentons ici des travaux non publiés qui ont mené à plusieurs perfectionnements de CLaP afin de solutionner certaines limites de la technique telle qu'exposée dans le chapitre précédent. Nous montrons tout d'abord comment CLaP peut être utilisé pour délivrer une molécule à la surface des cellules, puis nous modifions la technique pour attribuer trois étiquettes colorées simultanément. Enfin, nous illustrons le bon fonctionnement de CLaP en trois dimensions puis nous démontrons la faisabilité de trier des cellules en créant des adhésions artificielles avec le substrat. Nous discutons ensuite ces résultats et proposons une section matériel et méthodes associées à ces développements.

a) Utilisation de CLaP pour l'ancrage d'une tierce molécule à la membrane plasmique

Ainsi que nous l'avons démontré dans la précédente partie, CLaP permet d'attacher de la streptavidine à la membrane de cellules choisies. La streptavidine est une molécule tétravalente (qui peut lier jusqu'à 4 molécules de biotine) et par conséquent elle est un candidat de choix pour fixer un quelconque ligand à la membrane d'une cellule. En effet, si l'on couvre la membrane de la cellule de streptavidine en utilisant CLaP, de nombreux sites de liaison resteront libres et permettront théoriquement d'y lier toute autre molécule qu'on aura préalablement biotinylée. La biotinylation, soit le fait d'accrocher de manière covalente une biotine à une molécule cible, est une réaction simple, pour laquelle il existe un grand nombre de kits commerciaux. Une simple incubation de l'échantillon avec une solution du ligand biotinylé va permettre sa capture irréversible sur la membrane de la cellule fonctionnalisée par laser. La difficulté majeure de l'application de ce protocole résulte de la concentration très élevée de b4f utilisée pour CLaP. En effet, rincer efficacement une telle quantité de réactif sans affecter l'intégrité d'un échantillon biologique n'est pas aisé. C'est pourtant nécessaire puisque

toutes les molécules de b4f restées libres dans l'échantillon après le rinçage sont susceptibles de venir occuper les sites de liaison de la streptavidine par une réaction très rapide et irréversible.

Selon notre expérience, une méthode de choix pour rincer sans dommage un échantillon est de le plonger tout entier dans un très grand volume de PBS, et de recommencer jusqu'à trois fois l'opération. Comme démontré par la figure 7, ceci nous a permis de lier par CLaP de la biotine à la surface de cellules, puis, après avoir soigneusement rincé l'échantillon, d'y attacher de la neutravidine non fluorescente (la neutravidine est une streptavidine modifiée afin de réduire son attachement non spécifique à la membrane des cellules), pour finalement ajouter en troisième couche la protéine fluorescente mCherry, au préalable biotinylée. Cette manipulation simple s'est montrée déterminante dans l'application présentée au chapitre 6 de cette thèse. En effet, bien que dans l'expérience illustrée figure 7 la molécule administrée aux cellules soit une simple protéine fluorescente, beaucoup d'autres molécules effectrices pourraient ainsi être fixées à la membrane de la cellule pour, par exemple, délivrer des drogues directement à des cellules choisies.

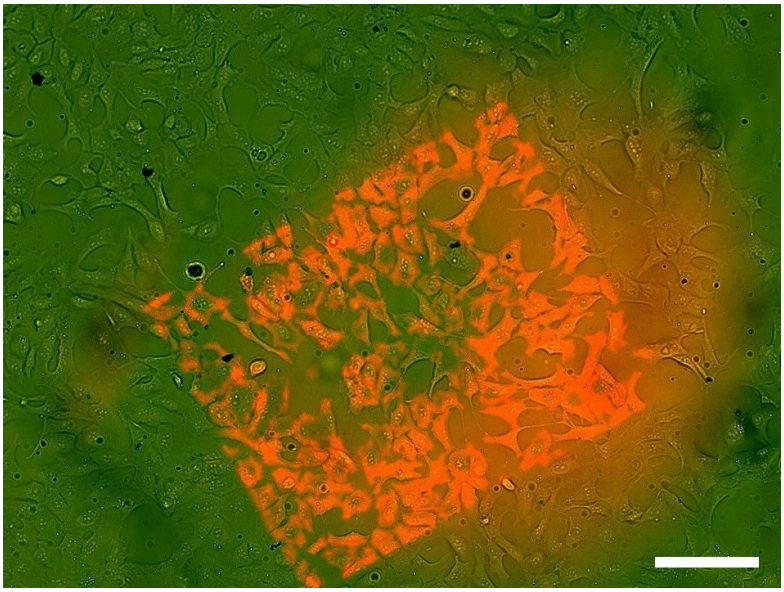


Figure 7. Utilisation de CLaP pour accrocher une molécule d'intérêt à la membrane des cellules cibles. Les membranes des cellules illuminées sont fonctionnalisées avec de la neutravidine non colorée avant l'addition d'une troisième couche moléculaire, ici, des molécules de mCherry biotinylées. L'image en fluorescence est superposée à l'image acquise en transmission afin de révéler tout l'échantillon, en particulier toutes les cellules qui n'ont pas été étiquetées. (Échelle : 100µm)

b) Marquage optique de cellules par des étiquettes multiples

Clap (*Cell Labeling via Photobleaching*) permet d'attacher une étiquette colorée à n'importe quelle cellule choisie arbitrairement dans une image de microscopie. Il devient dès lors possible de reconnaître des cellules, et ce, sans connaissance préalable d'un ligand membranaire spécifique de la cellule à marquer, ni d'un partenaire de réaction de forte affinité pour ce récepteur. Cependant, Clap utilise la biotine pour générer une étiquette, et une seule. Or, pouvoir identifier plusieurs cellules/types cellulaires simultanément ouvrirait les portes à de nombreuses autres applications. Une façon évidente de placer plusieurs tags sur différentes cellules dans un échantillon est de les ajouter séquentiellement un à un. Une première étiquette (par exemple une streptavidine rouge) est placée sur quelques cellules. Elle va occuper tous les sites de liaison (toutes les biotines), empêchant toute autre étiquette placée dans le milieu par la suite de s'y lier aussi. Une seconde marque peut dès lors être ajoutée en

répétant simplement le protocole complet (par exemple avec une streptavidine verte), etc… (Fig 8)

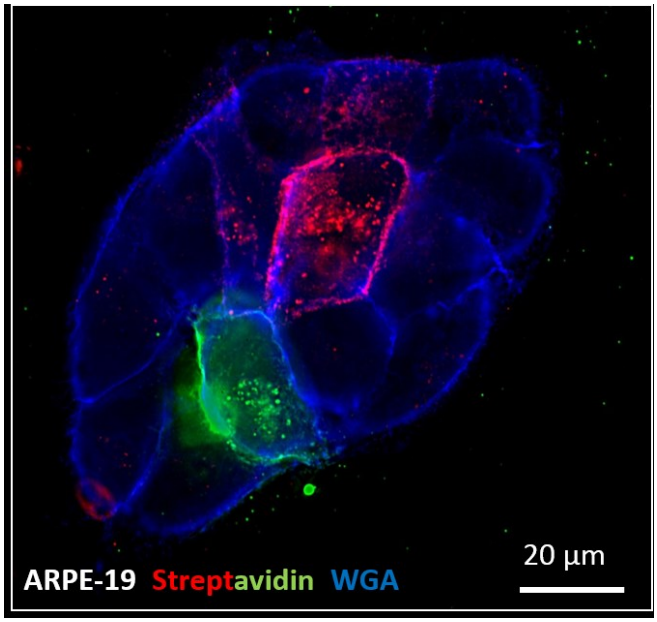


Figure 8. CLaP séquentiel en deux couleurs. Les membranes de cellules ARPE-19 sont colorées par incubation avec du WGA-alexa488 (ici en bleu) afin de rendre visibles toutes les cellules. Une première cellule est colorée avec une streptavidine-alexa 647 (en rouge) puis le protocole est entièrement répété afin de colorer sa voisine avec une streptavidine-alexa555 (en vert)

Cependant, lorsque le nombre de tags augmente le procédé devient rapidement laborieux et les trop nombreux rinçages et incubations augmentent la toxicité du protocole. Il est donc nécessaire de disposer de plusieurs étiquettes utilisables simultanément.

Puisque la liaison est créée par la formation d'un radical libre lors de l'excitation d'un fluorophore par un laser, il suffit d'utiliser différents fluorophores, chacun excité par une longueur d'onde qui lui est propre, pour pouvoir attacher des étiquettes moléculaires distinctes dans un même échantillon. Le choix du laser utilisé pour illuminer une cellule donnée détermine alors la marque qui est apposée sur sa membrane. Afin d'augmenter le nombre d'identités attribuables aux cellules, les couleurs peuvent être combinées. Ainsi, l'utilisation

de deux marques A et B génère 4 codes : aucune marque, seulement la marque A, seulement la marque B, et les marques A et B simultanément. Le nombre théorique d'objets identifiables est de 2^N, où N est ne nombre de fluorophores, ou 2^N-1 si l'on considère que l'absence de marque n'est caractéristique d'aucune cellule puisque c'est l'état dans lequel seront toutes les cellules non taguées.

Les concentrations très élevées de fluorophores nécessaires pour CLaP interdisent de fonctionnaliser directement la molécule cible avec le fluorophore, afin de maintenir un coût de réaction raisonnable. Dans l'exemple précédent, à la fois le fluorophore (fluorescéine) et la molécule d'intérêt (mCherry) étaient liés à une biotin et la streptavidine, tétravalente, permettait d'assurer leur dimérisation. Pour immobiliser une étiquette sur la membrane de cellules, on utilise une paire de petites molécules qui interagissent fortement et spécifiquement. L'une sera attachée au fluorophore, et l'autre sera liée à l'étiquette. Outre des fluorophores excitables par des longueurs d'ondes différentes, il est donc aussi nécessaire de disposer de plusieurs paires de petites molécules pour performer un CLaP dans plusieurs canaux. Ces petites molécules doivent être peu onéreuses, avoir une très grande affinité l'une pour l'autre, la réaction doit être rapide, aboutir à un produit stable, et doit être sélective : les molécules retenues ne doivent pas pouvoir se lier à un membre d'une autre des paires choisies, ni à des molécules naturellement présentes dans des échantillons biologiques.

Les réactions bio orthogonales fournissent de telles paires de molécules : l'acide azido-acétique se lie à la dibenzocyclooctyne (DIBO), la tétrazine se lie au transcyclooctène[153-154]. D'autres partenaires envisageables sont les SNAP-tags qui sont des séquences d'acides aminés qui se lient à O6-benzylguanine[155]. Un trait particulier de ce candidat est qu'il s'agit d'une séquence d'acides aminés, et qu'il peut donc être facilement inclus dans des protéines recombinantes. Une version modifiée, appelée CLIP-tag, a été produite et permet de sélectionner un autre partenaire d'interaction (O2-benzylcytosine) de manière orthogonale[155]. Enfin, de courtes séquences d'oligonucléotides peuvent fournir un très grand nombre de partenaires d'interaction ne réagissant pas entre eux. Dans la figure 9, nous avons utilisé trois paires de molécules : la classique paire biotine-streptavidine, l'acide azido-acétique qui lie la

dibenzocyclooctyne et enfin la tétrazine qui lie le transcyclooctène, ce qui porte à 3 le nombre d'étiquettes simultanément utilisables, et donc à 2^3=8 le nombre d'identités qui peuvent être attribuées dans un même échantillon en une seule incubation. Une section décrivant en détail la méthode permettant d'attacher ces trois identifiants cellulaires est fournie en annexe. Il est à noter que l'utilisation de trois marqueurs augmente significativement les temps d'incubation et de photoblanchiment, ce qui augmente simultanément le bruit de fond en laissant plus de temps aux cellules pour internaliser les différentes molécules fluorescentes. De plus, dans la version simple de CLaP, le canal utilisé pour observer les cellules (Alexa-647) est éloigné de celui utilisé pour le photoblanchiment (FITC, 488 nm) alors que l'utilisation d'un plus grand nombre de couleurs force l'utilisation des mêmes canaux pour fixer et observer le tag. Il reste aisé (Figure 9) de différencier sur une image le signal généré par CLAP (situé sur l'ensemble de la membrane des cellules) de celui résultant d'une internalisation non spécifique (localisé dans des vésicules dans la cellule). Cependant, la présence de ce signal non spécifique empêche une analyse par FACS. Certaines des molécules utilisées, comme le transcyclooctène, ne sont pas solubles dans l'eau. Elles doivent au préalable être dissoutes dans du DMSO, qui est connu pour perméabiliser les membranes cellulaires et accroit donc le signal non spécifique. L'utilisation de molécules passives plus grosses sur lesquelles seraient fixés nos marqueurs est une solution à envisager pour ralentir leur internalisation non spécifique. De même l'addition de groupes sulfo- permettrait de rendre le composé soluble dans l'eau et d'éviter l'utilisation de DMSO.

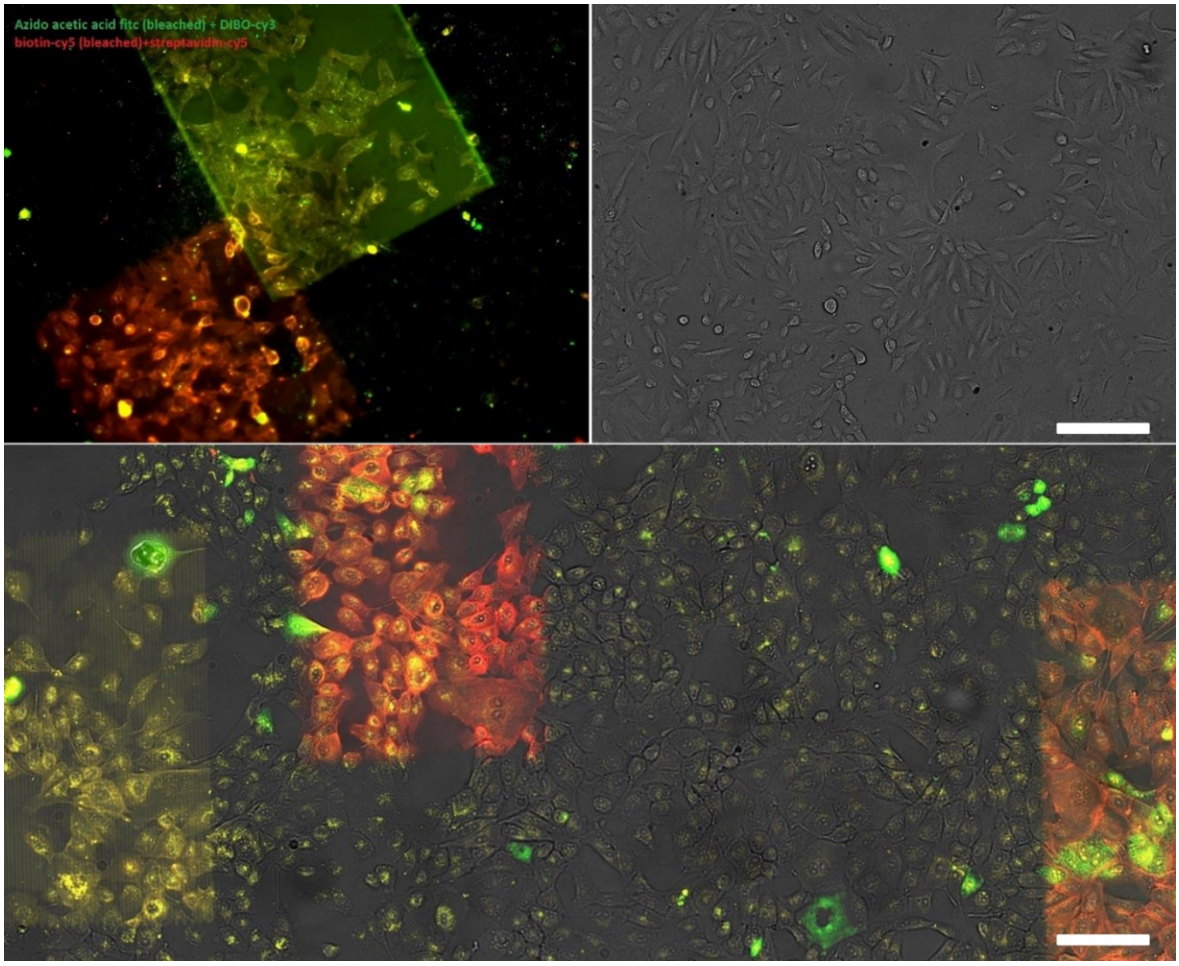


Figure 9. CLaP multicouleur. En haut à gauche, des cellules U2Os ont été taguées avec les paires azido acetic acid-DIBO (en vert) et biotine-strepatvidine (en rouge). A droite, une image en lumière blanche montre que ces cellules ont été spécifiquement taguées sans affecter les autres cellules de l'échantillon. En bas, une image sur laquelle trois régions différentes ont été taguées avec les paires biotin-streptavidine (en rouge), azido acetic acid-DIBO (en orange) et tétrazine-transcyclooctène (en jaune). Une image en transmission a été superposée afin que toutes les cellules soient visibles. Cette image en 3 couleurs, illustre bien le problème posé par l'internalisation non spécifique des colorants par les cellules. Échelle : 150µm.

c) Utilisation de CLaP dans un échantillon tridimensionnel

Les résultats présentés jusqu'ici montrent que CLaP est bien une solution au problème posé dans le cas de cultures cellulaires en deux dimensions. Cependant, l'utilisation de cette méthode dans la majorité des échantillons biologiques requiert qu'elle soit aussi fonctionnelle dans un environnement tridimensionnel.

Dans un premier temps, nous avons montré que la technique peut être utilisée *ex-vivo*, sur des explants de rétine de rat. Pour cela nous avons incubé ces explants dans de la biotin-4-fluorescéine pendant une heure afin qu'elle diffuse à l'intérieur du tissu puis nous avons illuminé un carré inscrit dans la rétine en utilisant le laser bleu d'un microscope confocal. Après rinçage, la rétine est incubée dans de la streptavidine colorée (ici, avec un alexa-647) pour révéler les cellules sur lesquelles a été attaché le b4f**.** La figure 10 montre une rétine de rat dont toutes les membranes cellulaires ont été colorées en vert en utilisant du WGA-alexa488. Plusieurs régions carrées ont été illuminées avec des intensités et temps d'illumination différents afin d'illustrer l'effet de la variation de ces paramètres. La figure montre un zoom sur la région la plus satisfaisante en termes de signal qui résulte d'une illumination de 20 secondes avec une puissance de 300 µW et un objectif d'ouverture numérique 0.4.

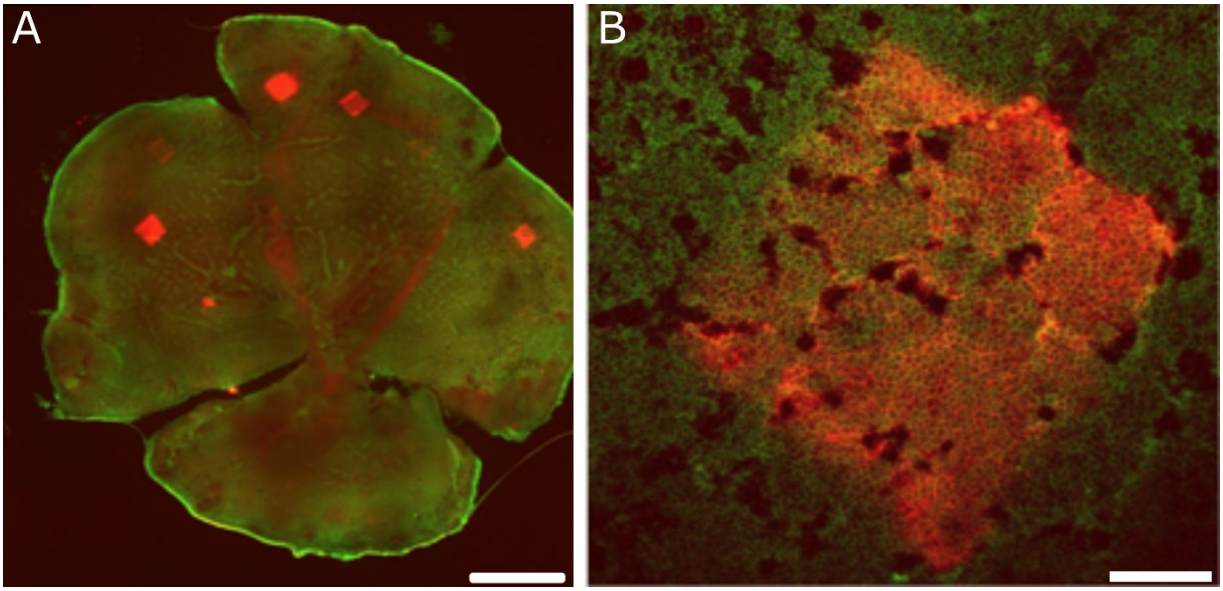


Figure 10. ClaP ex-vivo : marquage d'une rétine de rat. A. Une rétine de rat présentée en « flat mount », taguée avec du WGA-alexa488 pour révéler toutes les cellules. De petites carrés ont été taguées avec CLAP, en utilisant différentes puissances de laser. Échelle : 200µm. **B.** Un grossissement sur le carré tagué avec 300µW pendant 20 secondes. Échelle : 80µm.

Dans cette expérience, la longueur d'onde du laser correspond à la longueur d'onde d'excitation du fluorophore, et il s'agit donc d'une excitation ''classique'', en un photon. Ainsi, tout un cône de lumière autour du point focal contient suffisamment d'énergie pour photoblanchir le fluorophore (Figure 11 A). Cela aboutit à la fonctionnalisation de toute une colonne de tissu, avec une bonne résolution latérale, mais une faible résolution dans l'épaisseur puisque le faisceau laser traverse tout l'échantillon. Même si le photoblanchiment est moins efficace en dehors du point focal, il reste néanmoins suffisant pour taguer une épaisseur de plusieurs cellules dans l'échantillon. Afin d'obtenir une sélectivité selon l'axe vertical, il est nécessaire de remplacer cette excitation par une excitation deux photons. Brièvement, l'énergie portée par les photons qui excitent le fluorophore est divisée par deux (la longueur d'onde, elle, est donc multipliée par deux). La molécule doit toujours absorber exactement la même quantité d'énergie pour changer d'état quantique. Il devient donc nécessaire qu'elle absorbe simultanément deux photons de demi-énergie. Une absorption simultanée est un évènement extrêmement rare, qui est obtenu en utilisant un laser pulsé focalisé. La densité en photons dans le point focal au moment d'une pulsation du laser est telle que la probabilité d'absorption simultanée de deux photons n'est plus négligeable à cet endroit. Il devient dès lors possible d'exciter des fluorophores avec une bonne précision verticale, et donc d'étiqueter des cellules dans un environnement tridimensionnel sans nécessairement affecter simultanément toutes les couches supérieures et inférieures à la couche d'intérêt. Sur la Figure 11A, on voit la forme des volumes excités dans les deux cas : à gauche, avec une excitation classique, tout un cône est excité alors qu'à droite, en deux photons, seule un petit point fluorescent est visible (voir le zoom). L'excitation 1 photon est satisfaisante lorsque l'échantillon est 2-dimensionnel puisqu'il n'y a aucune autre couche à exciter (Figure 11A, extrême gauche) alors qu'une excitation 2 photons est requise quand l'échantillon est tri-dimensionnel, par exemple dans une rétine (Figure 11B, extrême droite). La figure 11 B montre trois vues sous différents angles d'un échantillon tridimensionnel dans lequel deux plans superposés ont été tagués. Cet échantillon est fait d'un gel de collagène dans lequel des cellules U2Os ont été cultivées. Une incubation longue permet au b4f de diffuser à

l'intérieur du gel, et deux plans superposés ont été photoblanchis pendant 20 secondes en utilisant un laser de 800 nm de longueur d'onde réglé à 150 µW. Après rinçage, une incubation avec de la streptavidine alexa-647 permet d'observer les plans biotinylés avec un microscope confocal. Des cellules choisies individuellement dans le gel tridimensionnel ont aussi été marquées (Figure 11 C). Pour cela, une excitation deux photons très faible est utilisée pour naviguer dans l'échantillon, les cellules apparaissant comme des objets noirs sur un fond saturé de fluorophore. Une région de la taille souhaitée est alors scannée avec une puissance plus élevée (100µW, 3 secondes) pour photoblanchir le b4f dans le voisinage immédiat de la cellule d'intérêt. Le b4f restant est ensuite rincé par immersions successives pendant une heure dans du milieu, puis l'échantillon est incubé dans de la streptavidine fluorescente (alexa-647) afin de révéler les cellules biotinylées. La streptavidine étant une molécule bien plus grosse que la biotine, l'élimination complète des molécules non liées nécessite un rinçage plus long, jusqu'à une nuit pour un gel épais. Un protocole est fourni en annexe pour détailler l'utilisation d'une excitation deux photons pour taguer des cellules par CLaP.

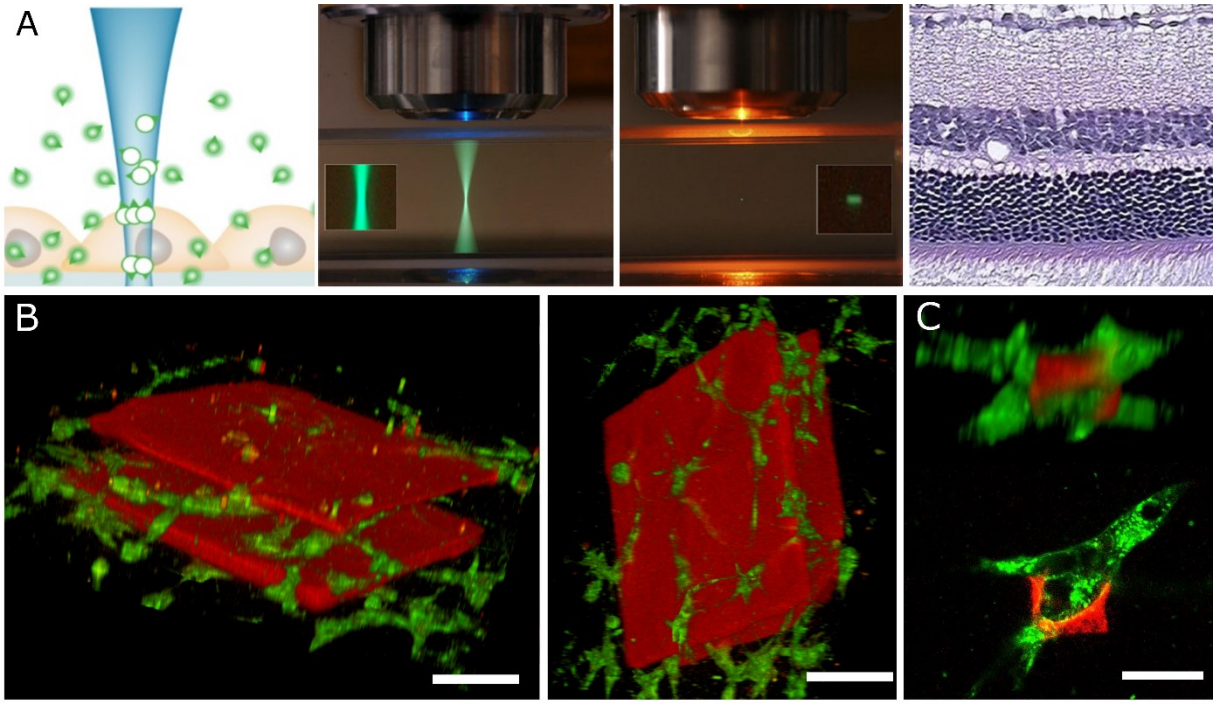


Figure 11. CLaP en trois dimensions. A. Comparaison des excitations à un photon (à gauche, en deux dimensions) telle qu'utilisée dans le protocole de CLAP et à deux photons

(à droite, en trois dimensions) telle qu'elle pourrait être utilisée dans un échantillon tridimensionnel comme celui représenté par la coupe histologique de rétine à l'extrême droite. **B.** Échantillon de collagène dans lequel des cellules U2Os (en vert, WGA-alexa488) ont été cultivées en trois dimensions. Deux plans superposés ont été tagués avec CLaP. **C.** Une cellule individuelle cultivée dans un gel de collagène a été taguée en deux photons. Toutes les cellules apparaissent en vert grâce à un marquage au WGA-alexa488.

Le choix de l'un ou l'autre des deux modes d'excitation permet de taguer soit toute une colonne de tissu, soit une cellule d'intérêt sans affecter les couches supérieures et inférieures. Les très longs temps de rinçage nécessaires pour permettre la diffusion de l'excès de réactifs hors de l'échantillon sont une limitation majeure de cette application de CLaP. En outre l'utilisation d'une excitation deux photons photoblanchit extrêmement efficacement les fluorophores et rendrait compliquée l'utilisation simultanée de plusieurs étiquettes associées à différents fluorophores.

d) Tri cellulaire par création de liaisons avec le substrat

Une dernière application que nous évoquons dans l'article *Live single cell laser tag* consiste à créer des liens entre les cellules marquées par CLaP et le substrat de culture pour trier ces cellules. En effet, lorsqu'un radical libre est capté par la membrane cellulaire, une réaction de polymérisation est initiée et peut mener à la création de liens artificiels entre la membrane plasmique et le support de culture. Dès lors qu'une cellule est artificiellement accrochée à son support de culture, il suffit de briser tous les liens naturels (par exemple en incubant l'échantillon dans de la trypsine ou de l'EDTA) pour enlever par simple rinçage toutes les autres cellules présentes dans la culture, et obtenir un échantillon qui ne contient plus que les quelques cellules soumises au protocole de CLaP.

Nous avons montré la faisabilité de ce tri en accrochant quelques cellules au plastique sur lequel elles étaient cultivées. Toutes les autres ont été décollées par une incubation dans de la trypsine, afin de ne garder en culture que ces cellules choisies arbitrairement. Dans nos mains, ce procédé s'est montré très efficace pour trier de petits groupes de cellules, tel qu'illustré figure 12, cependant, le tri d'une seule cellule, moins fortement liée à son substrat, s'est révélé plus difficile.

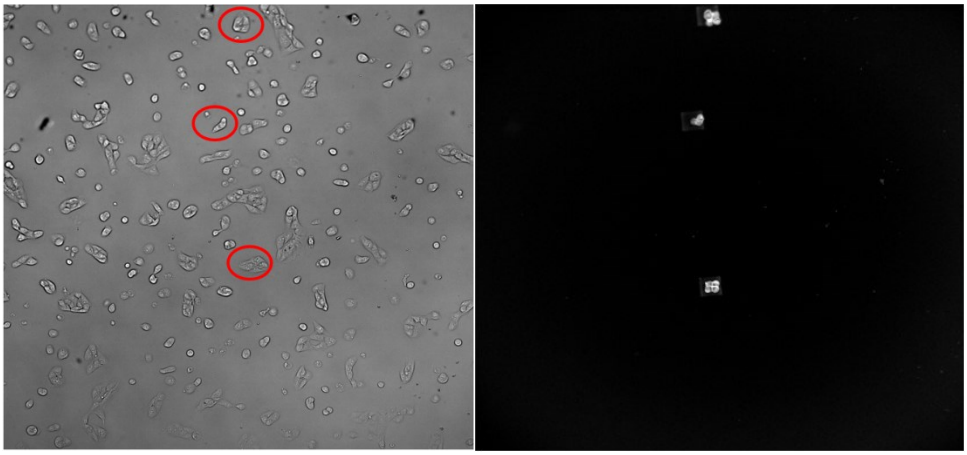


Figure 12. Tri cellulaire par attachement au substrat. A gauche, une culture deux dimensionnelle de cellules sur du plastique. Trois amas de cellules sont choisis pour être liés au substrat. À droite, après incubation dans de la trypsine, seuls ces trois amas sont encore attachés au substrat de culture.

e) Discussion

Nous avons proposé quelques solutions aux principales limitations de CLaP exprimées dans l'article pour permettre son utilisation dans des structures tridimensionnelles, des échantillons ex-vivo, et pour autoriser le placement simultané de plusieurs étiquettes. CLaP permet d'identifier efficacement des cellules individuelles selon des critères visibles en microscopie, sans connaissance préalable de ce qui fait leur particularité. Une limitation non abordée de la technique est son débit extrêmement bas dans sa forme actuelle. Cependant, l'identification des cellules d'intérêt dans des images ainsi que la quantification de leur mouvement dans l'échantillon peuvent être automatisés afin d'augmenter significativement le débit de la méthode.

Toutefois, comme expliqué dans l'introduction de ce travail, l'extraction et la purification des rares cellules taguées reste un défi. Dans de nombreuses applications, une fois les cellules d'intérêt reconnues, elles doivent être extraites pour être séquencées, cultivées en une nouvelle lignée de caractéristiques choisies, ou réinjectées dans un organisme. Les techniques courantes de tri cellulaire sont basées sur la micro fluidique afin de trier rapidement un grand nombre de cellules[17], ce qui est généralement nécessaire étant donné la taille des échantillons considérés. Cependant, ces précieuses cellules sont très souvent

perdues, ou capturées aux côtés de nombreuses cellules négatives contaminantes[156-157]. Les solutions basées sur des aimants n'offrent pas non plus une efficacité et une pureté satisfaisantes[158-159]. De plus, ces techniques utilisent encore l'interaction entre un marqueur typique de la cellule qu'il faut connaître a priori et un partenaire d'interaction spécifique de ce marqueur qui doit exister et être un réactif disponible. La technique présentée dans la partie d*)* précédente permet de détacher du substrat toutes les cellules d'une culture sauf certaines cellules choisies. Cependant, sa mise au point nécessite un travail de calibration très précis pour se placer dans les conditions qui permettent de détacher toutes les cellules négatives sans affecter les cellules sélectionnées. Ces paramètres dépendant du type cellulaire, de l'âge et de l'état de la culture. De plus, cette procédure souffre d'une mauvaise répétabilité. Aussi, nous avons développé une autre méthode pour trier des cellules avec une bien meilleure répétabilité qui est l'objet du chapitre suivant.

f) Matériel et méthodes

i) Culture cellulaire

Afin de réduire les quantités de réactifs nécessaires, nous utilisons de petites chambres de 5mm de diamètre coupées avec un emporte-pièce dans du PDMS et collées sur des lamelles de verre (le PDMS colle naturellement au verre). Le PDMS est préparé en coulant 10ml de monomère mélangé à 1ml d'agent de réticulation dans un pétri de 10 cm de diamètre. Le pétri est placé sous vide une nuit pour en retirer les bulles d'air. Le gel est prêt à être utilisé 48h après avoir été coulé.

Les cellules sont cultivées dans du DMEM/F12 additionné de 10% FBS et d'antibiotiques sauf lorsque précisé.

ii) Préparation d'échantillons 3-dimentionnels

Le collagène est préparé sur glace en ajoutant 20µl de milieu 10x à 180µl de collagène (Thermofisher, A1048301) et est amené à pH 7 en ajoutant du NaOH jusqu'à ce que le rouge

de phénol du milieu redevienne rose. Le pH peut être vérifié en utilisant du papier pH. 5µL d'une suspension concentrée de cellules U2Os sont alors ajoutés et mélangés avec le collagène. Le gel est pipetté dans un MatTek (pétri dont le fond est en verre) préalablement refroidi sur glace (permet d'éviter le décollement du gel lorsqu'il est plus tard immergé dans du milieu). Le gel est placé dans un incubateur à 37°C, 5% CO_2 pendant une heure, puis immergé dans 2ml de milieu (DMEM/F12, 10%FBS, antibiotiques) et gardé en culture 2 jours avant d'être utilisé pour CLaP.

iii) Protocole pour délivrer des protéines à la membrane de cellules choisies.

Le milieu (DMEM/F12 additionné de 10% FBS et d'antibiotiques) est remplacé par de la biotin-4-fluorescéine (0.05mg/ml) diluée dans du milieu de culture. Une région est scannée ligne par ligne avec un laser de longueur d'onde 488nm et de puissance 300 µW, en faisant bouger l'échantillon à une vitesse de 0.05mm/s, le laser restant fixe. L'espace entre les lignes est réglé à 5µm. L'échantillon est ensuite rincé 6 fois par immersion dans du PBS, puis incubé 5 min dans une solution de streptavidine à 0.05mg/ml dans du milieu de culture. La chambre de culture est alors rincée à nouveau trois fois puis incubée avec du milieu contenant 0.02mg/ml de mCherry biotinylée pendant 10 minutes. Après trois rinçages supplémentaires, l'échantillon peut être imagé pour vérifier l'attachement de la protéine sur les cellules cibles.

iv) Protocole pour l'attribution simultanée de trois marques cellulaires.

Les échantillons sont préparés la veille, dans des chambres maison coupées dans du PDMS et collées sur du verre. Le milieu est additionné d'azide-alexa488, biotine-cy5 et tetrazine-cy3 pour atteindre une concentration de 60 µM pour chaque composé. Trois régions sont alors illuminées avec des puissances de 3mW en bleu (488nm), 5mW en vert (555nm) et 8 mW en rouge (647nm) à des vitesses de 0.04mm/s pour le bleu et le rouge, et 0.08mm/s pour le vert afin de photoblanchir spécifiquement chaque fluorophore. Après6 rinçages, l'échantillon est incubé avec 60 µM de dibenzocyclooctyne-alexa488, 1 µM de streptavidine-cy5 et 60 µM de trancyclooctene-cy3.

v) Utilisation d'une excitation 2-photons pour taguer des cellules en trois dimensions

Le gel contenant les cellules en trois dimensions est incubé pendant une heure dans 2 ml de milieu contenant 0.04 mg/ml de biotine-4-fluorescéine. Lorsque l'objectif est de taguer une région de l'échantillon, celle-ci peut être choisie par observation du gel en lumière blanche avant d'être scannée en utilisant les galvanomètres du système pendant 20 secondes avec une puissance de 150µW. Au contraire, lorsque l'objectif est de taguer des cellules individuelles, l'échantillon doit être observé en 2 photons avec une puissance minimale, afin de choisir une cellule d'intérêt et de délimiter la région et le plan à scanner (une région contenant la cellule d'intérêt seulement). La puissance du laser est alors ajustée à 100µW et la cellule est scannée pendant 3 trois secondes. L'échantillon est alors rincé pendant 45 minutes, puis incubé dans 2 ml de milieu additionné de 0.05mg/ml de streptavidine-Alexa-647 pendant une heure. Le gel est rincé à nouveau dans du milieu pendant 2 heures à une nuit en changeant le milieu plusieurs fois. (Au minimum une fois après 15 minutes, puis à nouveau après 30 minutes et une heure.) L'échantillon peut être incubé 20 minutes dans du WGA-alexa488 afin de colorer les membranes de toutes les cellules présentes dans le gel.

vi) Utilisation de CLaP pour isoler des cellules en les attachant à leur substrat de culture

Des échantillons de cellules à faible confluence sont préparés la veille. Les cellules (ou petits amas de cellules) sont taguées selon le protocole classique de CLaP. Pour cela, une région contenant les cellules d'intérêt est scannée avec des galvanomètres et un laser réglé à 488 nm et une puissance de 500 µW pendant 20 secondes. Les cellules négatives sont alors détachées avec 0.25% de trypsine ou de l'EDTA (10mM). La durée de l'incubation est à optimiser pour chaque type cellulaire : une incubation trop courte ne détachera pas toutes les cellules négatives, mais une incubation trop longue détachera toutes les cellules, y compris les cellules taguées avec CLaP. Généralement, 5 à 10 minutes permettent de conserver les cellules d'intérêt tout en décrochant toutes les cellules négatives. L'ajout de streptavidine fluorescente permet de vérifier l'identité des cellules conservées puisqu'elles devraient toutes être ainsi colorées.

Chapitre 6 : Enrichissement opto-magnétique de cultures cellulaires : article : *Opto-magnetic capture of individual cells based on visual phenotypes*

Nous avons adapté CLaP pour attacher des billes magnétiques aux cellules d'intérêt. Le caractère arbitraire de la ligation est conservé, et la sélectivité de la méthode rend possible l'identification de cellules une à une afin de les extraire avec une grande efficacité et une grande pureté. Les cellules triées sont viables comme nous l'illustrons en triant des cellules très sensibles comme des cellules primaires ou même des cellules souches embryonnaires. Nous avons généré une lignée cellulaire qui présente la particularité de réparer des dommages induits à l'ADN plus rapidement que la population cellulaire d'origine. Nous avons aussi purifié des cellules multinucléées, impliquées dans le développement des résistances aux médicaments et dans les récidives du cancer. Finalement, nous fournissons un troisième exemple d'utilisation de cette méthode en triant des cellules en cours de différenciation en adipocytes. Ce chapitre est constitué du texte tel qu'il a été accepté pour publication dans *eLife* en mai 2019. Les auteurs sont : Loïc Binan, François Bélanger, Maxime Uriarte, Jean François Lemay, Jean Christophe Pelletier De Koninck, Joannie Roy, Bachir el Affar, Elliot Drobetski, Hugo Wurtele, Santiago Costantino.

a) Summary

A new method, termed Single-Cell Magneto-Optical Capture (scMOCa), uses laser illumination to efficiently capture individual cells from large populations based solely on visual

characteristics observable in microscopy images. This technique, requiring only simple reagents, permits clonal expansion of chosen cells for subsequent in-depth analyses.

b) Abstract

The ability to isolate rare live cells within a heterogeneous population based solely on visual criteria remains technically challenging, due largely to limitations imposed by existing sorting technologies. Here we present a new method that permits labeling cells of interest by attaching streptavidin-coated magnetic beads to their membranes using the lasers of a confocal microscope. A simple magnet allows highly-specific isolation of the labeled cells, which then remain viable and proliferate normally. As proof of principle, we tagged, isolated, and expanded individual cells based on three biologically-relevant visual characteristics: i) presence of multiple nuclei, ii) accumulation of lipid vesicles, and iii) ability to resolve ionizing radiation-induced DNA damage foci. Our method constitutes a rapid, efficient, and cost-effective approach for isolation and subsequent characterization of rare cells based on observable traits such as movement, shape, or location, which in turn can generate novel mechanistic insights into important biological processes.

c) Introduction

Characterization of biological samples relies heavily on microscopy where, in response to various stimuli, molecular probes and a myriad of contrast reagents are routinely used to identify and label individual live cells of interest. These methods often require prior knowledge of cellular markers or use of elaborate reporter constructs. On the other hand, based solely on visual inspection or using image processing algorithms, it is possible to distinguish rare cells which exhibit distinct biological properties from among thousands of counterparts within a microscopy field. Such visually-discernable traits include movement, shape, intracellular protein distribution, and location within the sample, and in turn can reflect important

physiological features of individual cells. For example, cell migration (movement) is an essential determinant in normal embryonic development, wound healing, immune responses, tumor progression, and vascular disease[11]. Moreover, changes in cellular morphology (shape) constitute biomarkers of cellular growth, division, death, and differentiation, as well as of tissue morphogenesis and disease[12]. Cell-to-cell contacts (location) or distance to sources of chemical cues such as senescent cells, inflammation or necrotic tissue are critical factors in chemokinesis, differentiation, neural function, and immune responses[9]. Finally, expression and visualization of fluorescent fusion proteins permits the identification of cells presenting molecular behaviors of interest, such as differential relocalization of proteins to subcellular compartments or structures upon various stimuli. Unfortunately, however, isolation and expansion of single cells characterized by such easily-observable features is technically challenging, and indeed has not been accomplished to date.

We recently developed a method termed Cell Labeling via Photobleaching (CLaP)[46] allowing the arbitrary tagging of individual cells among a heterogeneous population within a microscopy field. This is accomplished by crosslinking biotin molecules to their plasma membranes with the lasers of a confocal microscope, followed by use of fluorescent streptavidin conjugates to reveal the marked cells. In this manner, the same instrument used for imaging can also be adapted to label particular cells based on any visible trait that distinguishes them from the ensemble. Importantly, previous knowledge of surface markers or transfection of reporter genes are not required. Tags can be added with single cell precision and the incorporated label displays convenient tracking properties to monitor location and movement. The mark is stable, non-toxic, retained in cells for several days, and moreover does not engender detectable changes in cell morphology, viability, or proliferative capacity. Moreover, gene expression profiling indicated no major changes associated with the procedure[46]. Nevertheless, a technology for the efficient isolation and expansion of CLaP-tagged cells is still lacking.

The fact that cell populations are often highly heterogeneous underscores the need for new approaches to capture and clonally expand individual cells of interest for further

characterization. However, as mentioned above, current sorting techniques cannot efficiently isolate such rare cells[17]; indeed classical protocols like Fluorescence and Magnetic Activated Cell Sorting (FACS and MACS) are typically optimized for high throughput at the expense of capture efficiency and specificity, and require large numbers of cells[17]. Small cell populations representing 10^{-3} of the total, which have been defined as rare, or ultrarare if their proportion is less than 10^{-5}, can only be effectively captured and purified with repeated cycles of sorting and cell expansion protocols[17]. Starting with rare and hence precious cell populations, highly conservative gating strategies are needed, which can at best achieve approximately 45% purity[156-157]. Time-consuming manipulations, cost, hardware footprint, and handling complexity[160] make approaches based on microfluidics ill-suited for capturing small numbers of cells, which are often masked within tens of thousands.

Here we report a novel technology, termed Single-Cell Magneto-Optical Capture (scMOCa), for isolating cells based purely on visual traits from within large heterogenous populations. After tethering biotin moieties to their membranes, cells of interest are targeted with streptavidin-coated ferromagnetic beads and captured with high efficiency using a simple magnet. The procedure is fast, uses low-cost commercially-available reagents and only requires access to a standard confocal microscope. As proof-of-principle for the utility and power of this novel approach, we used scMOCa to i): capture and expand individual cells that differ in their capacity to resolve ionizing radiation (IR)-induced foci of the DNA repair protein 53BP1, ii) purify rare multinucleated cells, and iii) isolate cells that differentiated into adipocytes and accumulated lipid vesicles. Overall, the ease of use and affordability of our method is expected to facilitate the characterization of phenotypes of interest occurring in a small fraction of cell populations.

d) Results

i) scMOCa: efficient magnetic sorting of cells using ferromagnetic streptavidin-coated beads

Cell membrane biotinylation and ferromagnetic functionalization

We set out to evaluate whether individual cells illuminated with a low-power laser can be labeled with ferromagnetic beads, thereby facilitating their purification and clonal expansion. Adherent cells were incubated in medium supplemented with biotin-4-fluorescein (B4F), and a small area inside the cells of interest was illuminated with a 473nm excitation laser at low power (<100 µW) for two seconds using a confocal microscope. This operation effectively crosslinks biotin molecules to plasma membranes and was repeated for all targeted cells. After washing, streptavidin-coated ferromagnetic beads were added to the medium, and then allowed to settle and attach specifically to illuminated cells (Figure 13A).

The high strength of the biotin streptavidin bond ($K_d=10^{-15}M$) allows stringent rinsing and efficient removal of unbound magnetic beads, which is key to obtaining specific tagging allowed by the accurate laser pointing (Figure 13B). Depending on their size, beads may later be internalized (nanometer-size beads) or retained at the cell surface and shared between daughter cells after mitosis (micron-size beads). If needed, special beads, which integrate a DNA spacer between the streptavidin and their magnetic core to allow enzymatic cleavage, are commercially available (Supplementary figure 13). This permits detachment from cells in cases where beads can compromise downstream experiments, e.g. analysis of migration, or single-cell RNA sequencing.

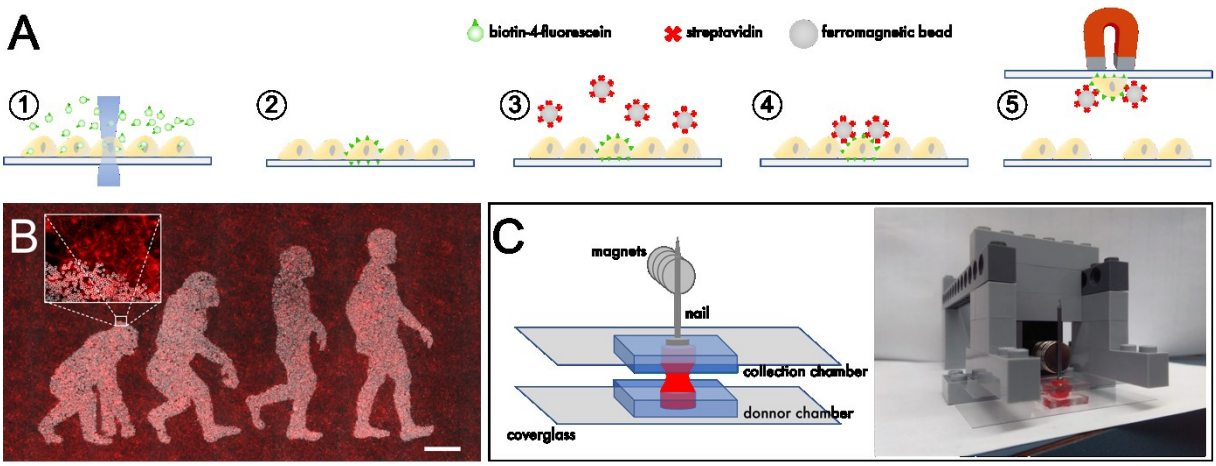


Figure 13. Description of the technique. A. Outline of scMOCa. Biotin-4-fluorescein is crosslinked to cell membranes with a laser. Biotin-tagged cells are labeled with streptavidin-coated ferromagnetic beads and captured with a magnet. **B.** Example of a confluent U2OS cell culture where only cells illuminated with the lasers of a confocal microscope are densely decorated with magnetic particles. Beads appear in white, and all cellular membranes in red, tagged with WGA-Alexa647. Scale bar: 500µm. **C.** Schematic illustrating the simple tools needed to implement the protocol. Two small cell culture chambers cast in silicone and adhered to coverglasses are positioned one on top of the other. Cells in the bottom chamber are attracted to the top collection chamber by a magnetic field. A nail is placed above the collection chamber to guide the field generated by magnets to the donor chamber in which the cell suspension is kept. The collection chamber is held between two Lego bricks, filled with a solution of Trypsin (held in place by surface tension), and then slowly approached 6 mm above the bottom chamber, at which point the two drops merge.

Rare cells can be sorted and expanded with high efficiency and specificity

We used trypsin to detach cells from the substrate before subjecting the entire population to a magnetic field that attracts labelled (positive) cells upwards to a collection chamber, while non-labeled (negative) cells remain in the original chamber. Specifically, two home-made chambers cast with silicone were filled with cell culture medium and positioned one on top of the other (Figure 13C). The top (receiving) chamber is also filled with trypsin and slowly brought together with the bottom chamber until both liquid drops merge. On top of the receiving chamber, a nail is placed to guide the magnetic field generated by a pile of ten N35 magnets, each generating a 1.18 Gauss magnetic field at its surface (Figure 13C). Importantly, the nail must have high iron-alloy content for strong ferromagnetism. Only positive cells coated

with ferromagnetic beads are pulled upwards to the top chamber, whereas negative cells are held down by gravity.

Magnets only attract positive cells with beads from the bottom well to the top well, regardless of the total number of cells in the sample. Repetition of the magnetic capture up to four times yields optimal selectivity: the collection (top) chamber can be simply flipped to replace the original donor chamber, while a new clean collection chamber is placed on top. The entire procedure takes only a few minutes and a detailed protocol is provided in materials and methods and supplementary figure 14. We note that a number of experimental parameters from this protocol need to be fine-tuned for specific cell types which exhibit different binding strengths and adhesion kinetics. In particular, the duration of the trypsin incubation, the number of times the capture is repeated, the time of exposure to the magnetic field, and the concentration of beads must be experimentally optimised.

Chamber dimensions can be critical for effective sorting, as their diameter (5 mm) and thickness (2 mm) determine the surface tension that holds liquid in the collection chamber and prevents it from falling. Furthermore, turbulence and movement must be avoided to prevent negative cells from reaching the collection chamber when both chambers are pulled apart. The distance that separates the two chambers while cells are being magnetically transferred must be maintained at approximately 6 millimeters such that gravity attracts negative cells as far away as possible from the collection chamber. The more distant the chambers are, the stronger the magnetic field must be to attract positive cells into the collection chamber; however, this could in turn affect the viability of transferred cells subjected to high pressure from beads pushing towards their cytoplasm.

We quantified the capacity of scMOCa to tag and isolate single cells from large populations. For this, we illuminated individual cells from chambers where approximately 50,000 cells had been seeded the day before and assessed capture efficiency. Figure 14 shows examples where 1 or 5 cells were successfully sorted. Cells were non-specifically stained with WGA-Alexa-555 to facilitate detection and images were obtained before (Figure 14, left panels) and immediately after sorting (Figure 14, right panels). The right panels of Figure 14 display

both captured cells (visible in red) as well as unbound beads often aligned with the magnetic lines of force emanating from the head of the nail.

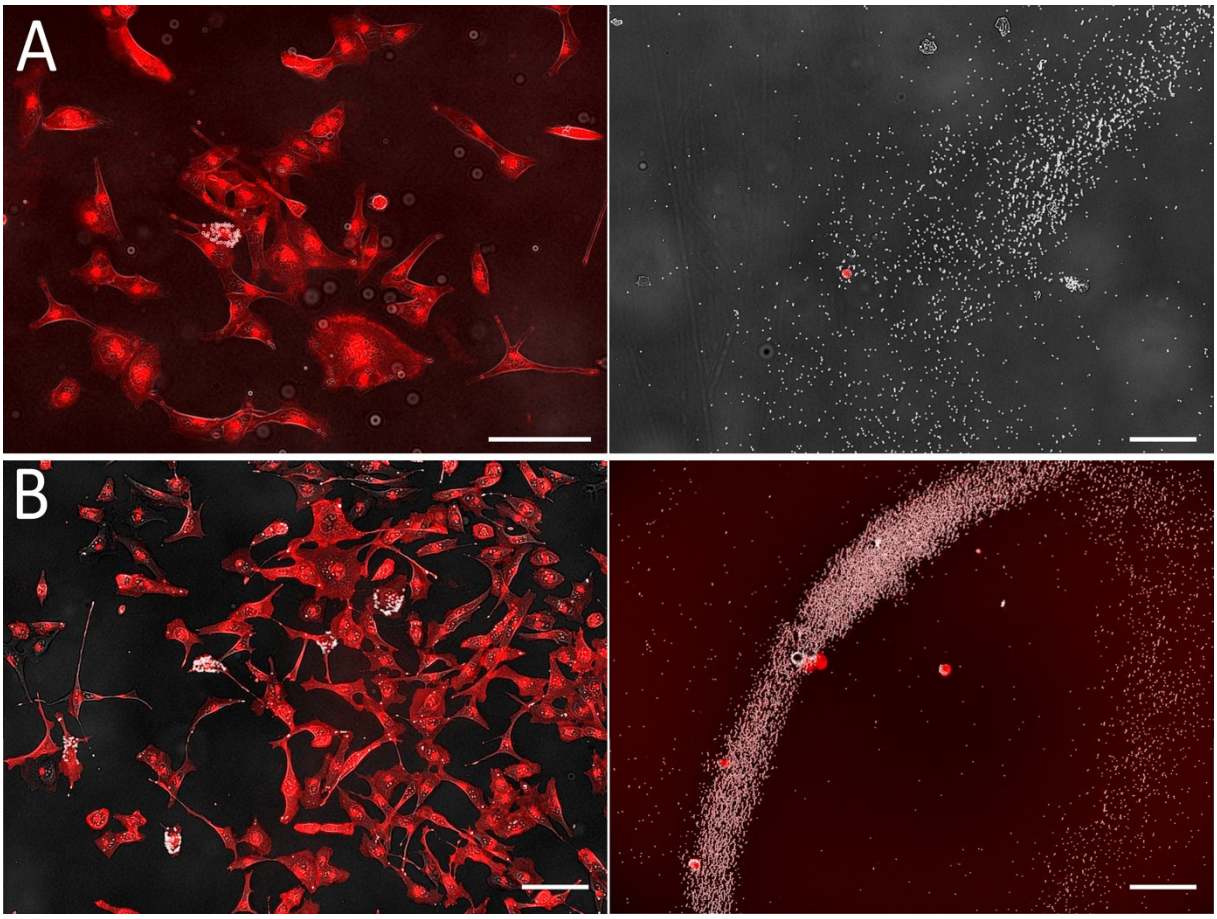


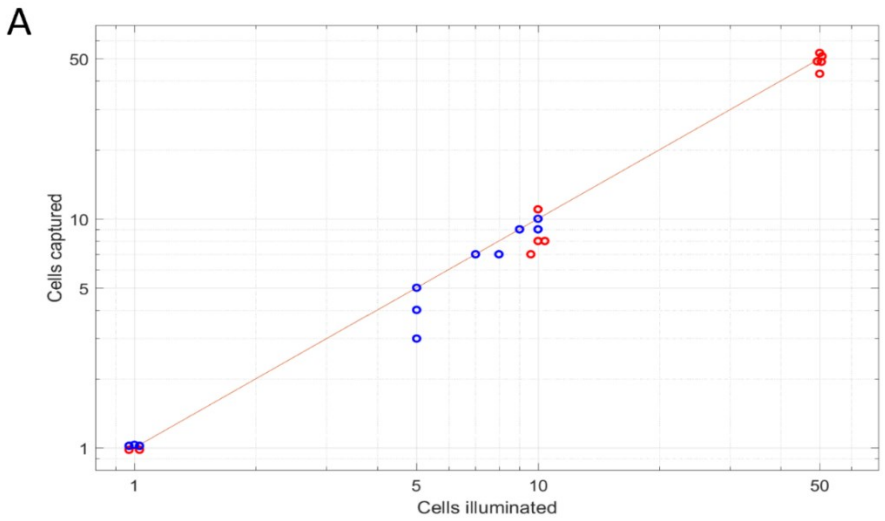
Figure 14. **Images of cells functionalized with magnetic beads** before (left, original chamber) and after (right, collection chamber) sorting. Beads appear in white (transmission image), and plasma membranes, tagged with WGA-Alexa555, in red (fluorescence image). Experiments were performed by tagging and sorting 1 cell **(A)** or 5 cells **(B)**. In each case, it is apparent that all selected cells (left) are efficiently extracted (right) without contamination as the number of cells on the images on the right corresponds to the number of cells tagged. Tagged cells are easily recognized as they are covered with beads in both images. Scale bars: 50μm.

We have repeated these experiments and obtained similar results using both glass and Aclar (plastic) substrates, which vary significantly in their ability to promote cell adhesion. In every experiment a given number of fluorescently labeled U2OS cells (1 to 50) were illuminated with a laser, sorted, and the receiving chamber examined to count captured cells. Cells in the receiving chamber with no visible beads attached to their membrane were considered as negative captured cells. Figure 15A demonstrates the high capture efficiency and selectivity of scMOCa, where blue dots correspond to experiments performed on Aclar substrates (higher cell adhesion) and red dots to glass (lower cell adhesion). Out of 23 experiments, starting from samples of 50,000 cells, the largest deviation from perfect recovery corresponds to one test where only 3, instead of 5 positive cells, were captured (2 positive cells lost).

To further demonstrate the high specificity of our capture technique, i.e. to determine the ratio of false positive cells to the total number of chosen cells, 50 000 cells originating from two different species were co-cultured: MDCK (dog kidney cells) and IMCD (mouse kidney cells) at a 1:1 ratio. IMCD cells were incubated in WGA-Alexa 555 prior to mixing, to add a species-specific fluorescent marker. After one day in co-culture, the sample was brought to the microscope where 10 (non-fluorescent) MDCK cells were illuminated. We sorted the cells using scMOCa and performed PCR with primers specific for the cytochrome C gene from both dog and mouse. The results show that both cell types were present in the original mix, but only dog DNA was detected after magnetic sorting (Figure 15B). We also show by qPCR that these samples respectively contain an amount of DNA that corresponds to 10 and 9 dog cells, whereas mouse DNA is essentially undetectable (Figure 15C). We also note that since we amplified a mitochondrial gene present at hundreds of gene copies per cell, one negative cell or even a DNA dilution corresponding to less than one cell is expected to be detectable (DNA dilutions corresponding to less than one cell give readily detectable signals; see calibration curves in supplementary figure 17). These experiments demonstrate that scMOCa isolates individual cells with high specificity. Indeed within a heterogeneous population, i.e. starting with a ratio 1:10,000 (positive: negative cells) in the source chamber, the method yields pure samples in the collection chamber. Our examples represent a five-orders-of-magnitude

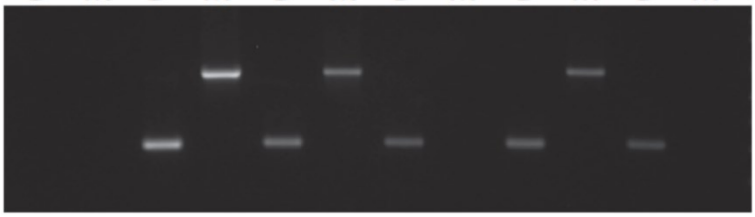
enrichment, as pure samples originating from a rare cell population (0.02% of the total) can be generated.

As a comparison to other capture methods based on magnetic fields, we prepared samples in which we sought to isolate 30 U2OS cells arbitrarily tagged amongst 30 000 by using commercially available separation columns (MACS, Miltenyi Biotec). These columns are optimised for high throughput enrichment of large samples and are not designed for rare cells. In three independent experiments we could isolate 5.3 ± 1.5 positive cells on average, while also capturing 17.6 ± 7.3 negative cells. This represents a population in which approximately 75% of the captured cells are contaminating false-positive cells with no beads attached, while scMOCa generates pure samples (Fig. 15A). These results underscore the importance of the design of the home-made chambers and capture protocol, which prevents turbulent movement of cells.



B

	no DNA		control DNA		Experiment A				Experiment B			
					Total		Picked		Total		Picked	
PCR:	D	M	D	M	D	M	D	M	D	M	D	M



C

	Experiment A		Experiment B	
	Dog	Mouse	Dog	Mouse
Percentage of cells in original sample	57 ± 6	43 ± 13	59 ± 8	41 ± 3
Number of cells picked	9.5 ± 0.6	< 0.01	8.9 ± 0.2	< 0.01

Figure 15. Evaluation of scMOCA effisciency. A. Capture efficiency for 1, 5, 10, and 50 selected cells for a total of 27 experiments. Red dots represent experiments performed with glass as a cell culture substrate and blue dots correspond to experiments using Aclar as a substrate. The horizontal axis represents the number of target cells, considered as the number of cells illuminated with the laser. Ordinate axis shows the number of cells detected on the collection chamber after capture, and the line corresponds to 100% success rate. A linear fit of the data yielded a slope of 0.99, demonstrating that scMOCa is highly efficient in retrieving all target cells, after testing 1 to 50 cells. **B, C.** Mouse (fluorescent) and dog (non-fluorescent) cell lines were co-cultured and only dog cells were illuminated and captured. PCR on a mitochondrial gene shows that all extracted cells form a pure sample and are exclusively dog cells. Table C shows the number of cells detected in each condition in three repeats of the experiment. These numbers are calculated from the amount of detected DNA normalized to the expected amount in one cell. A and B are independent experiments in which two different dishes were prepared, tagged and sorted prior to PCR.

Cells can be placed back in culture and expanded after sorting. Immediately after capture cells are round (as expected after trypsin treatment), but after one day in culture they display normal elongated shapes (Figure 16). Upon proliferation the number of cells with beads attached is reduced exponentially as cells divide (Figure 16, right panels). In addition to immortalized cell lines, we have tested and successfully sorted three different types of primary cells: human umbilical vein endothelial cells (HUVECs), human lung fibroblasts, and mice dorsal root ganglion (DRG) neurons dissected and plated 24 hours before the assay. We specifically chose primary cells as these are known to be more fragile during manipulation than cell lines. Importantly, HUVECs and lung fibroblasts proliferated normally for several days and primary DRG neurons actively extended cellular processes, as shown in Figure 16. Finally, we tested mouse embryonic stem cells which, after capture and replating, displayed similar growth and morphological features relative to the original population. Indeed, cells sorted using gelatin coated plastic chambers migrated and regrouped into small colonies which proliferated normally during 10 days. Sorted cells formed small poorly adherent spherical structures (figure 16D) which is expected from embryonic stem cells as they are known to spontaneously form embryonic bodies in culture. Upon addition of 1uM retinoic acid and removal of the leukemia inhibitory factor (LIF) from their medium, they started differentiating during 5 additional days (figure 16D, right panel) and became more adherent cells spread on the culture substrate.

High plating efficiency is important when only one sample with very few cells needs to be expanded. Therefore, chamber culture conditions must be optimized for low cell numbers. Cell viability and proliferative potential can be improved by the use of conditioned medium[161-163], i.e., medium collected from an exponentially growing cell culture and passed through a 0.2µm filter. This is attributed to secreted factors that in turn facilitate cell growth at very low density[161, 163]. The top collection chamber can be coated with collagen to further improve cell attachment and viability[164].

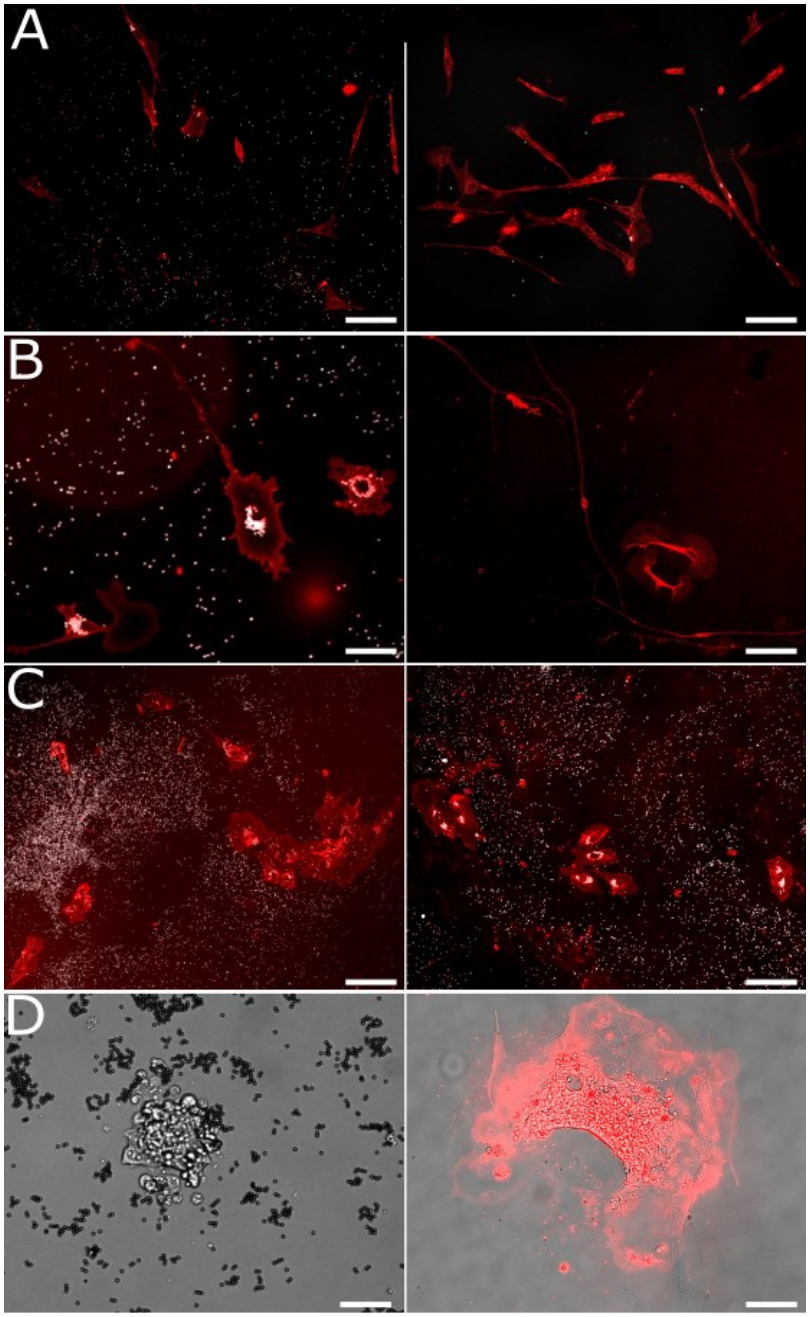


Figure 16. Primary cells and embryonic stem cells survive scMOCA. A. Images showing scMOCa-captured cells stained with WGA-Alexa-647. **A.** LF-1 fibroblasts 1 (left) and 4 (right) days after sorting. Scale bar 80µm. **B.** Primary DRG neurons 2 (left) and 4 right) days after sorting. Scale bars: 25µm (left) and 80µm (right) **C.** HUVECs 3 (left) and 6 (right) days after sorting. Scale bar: 80µm. **D.** Mouse embryonic stem cells seven days after sorting (left) and 5 days after starting differentiation (15 days after sorting) (right). Prior to differentiation, only a bright field image is shown to preserve cell viability. After differentiation, we stained cells with WGA-Alexa647, and merged the image with a bright field photo to increase contrast and better see cellular extensions. Scale bar: 40 µm.

ii) Cells can be captured based on their ability to resolve ionizing radiation-induced DNA damage foci.

To demonstrate the utility of scMOCa, we sought to isolate and expand cell populations based on their ability to resolve ionizing radiation (IR)-induced 53BP1 DNA damage foci, a well-characterized indicator of DNA double strand break (DSB) repair capacity[165]. For this, we used U2OS osteosarcoma cells harboring a construct permitting doxycycline-inducible expression of 53BP1 fused to Green Fluorescent Protein (GFP). 53BP1 is directly involved in DSB repair and is rapidly recruited to DSB sites where it forms foci that can be readily detected by fluorescence microscopy in live-cells[166] when fused with GFP. Foci of 53BP1 are resolved gradually as cells repair DSB, and within approximately 3 hours post-irradiation with 0.5 Gy most are expected to disappear[166].

We exposed cells to 0.5 Gy of IR and imaged GFP-53BP1 foci. We first characterized focus formation and resolution by measuring the average number of foci before and after IR in 500 cells. At 45 minutes post-irradiation an average of 10.2$\pm$2.5 (mean $\pm$ standard deviation) foci per cell was detected. At two hours post-irradiation, a second set of images was acquired, and the average number of foci was reduced to 7.6$\pm$2.3. Since on average cells resolved approximately 25% of their foci within two hours, we defined cells in which more than 85% of foci have disappeared after two hours as “fast resolving”. Such fast resolving cells, represented approximately 1% of the population. In all following experiments, we compared both sets of images to search for fast-resolving cells (two such cells are shown in Figure 17A) and used scMOCa to tag, capture and expand them.

We emphasize that FACS or similar approaches are not suitable for sorting based on focus resolution, even if the fraction of target cells was relatively large, as the overall fluorescence signal originating from cell nuclei does not reflect the local distribution of protein. Indeed, we observed no change in global protein abundance or average intensity of GFP-53BP1 upon focus resolution: the average intensity of nuclei showed no correlation with the number of 53BP1 foci (Pearson coefficient of -0.15). Because we used very stringent selection criteria for focus resolution, we tagged only 5 and 3 “fast-resolving” cells in two independent

experiments, which were subsequently isolated using scMOCa, pooled and expanded to generate Populations #1 and #2.

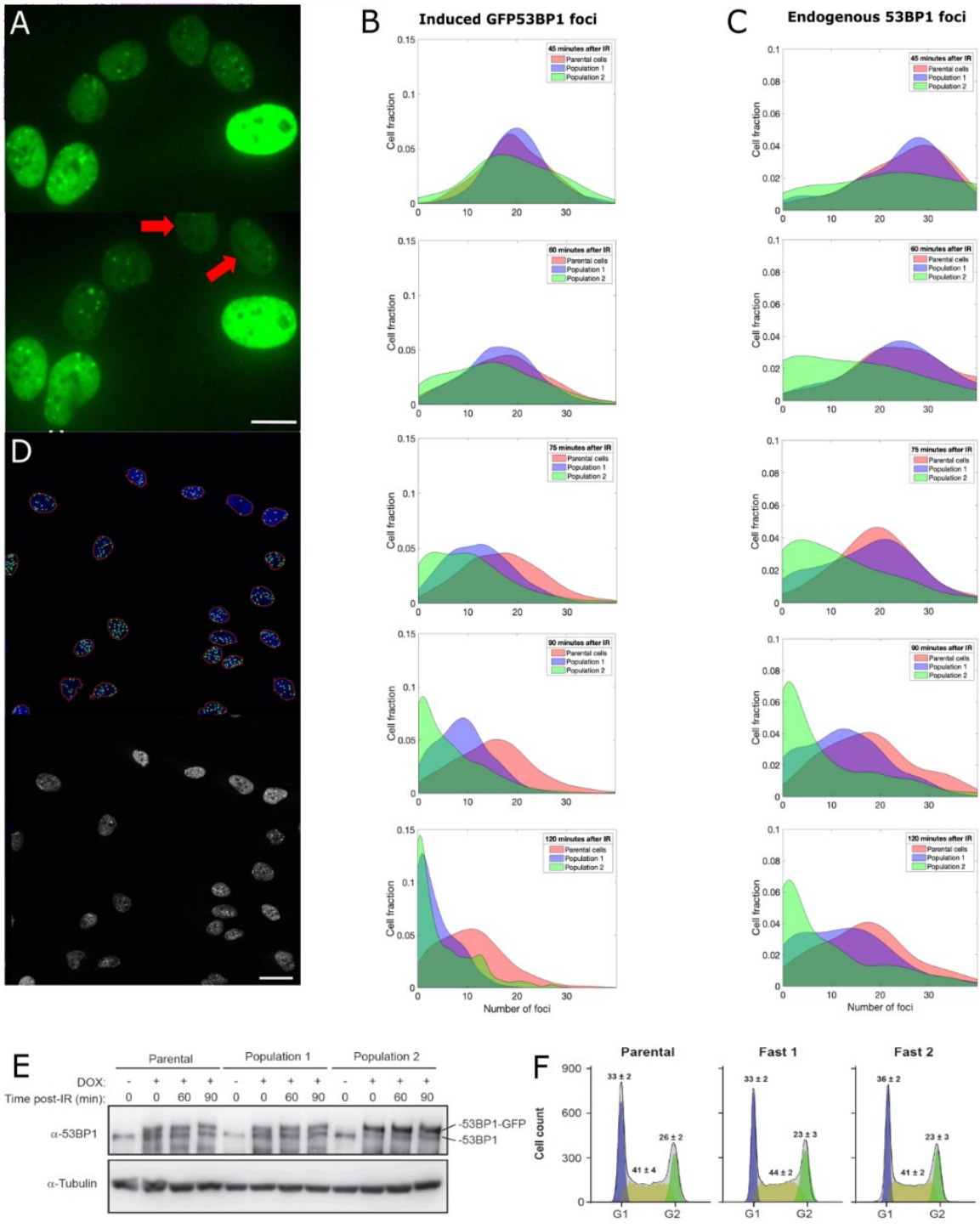


Figure 17. Capture and expansion of individual cells that differ in their capacity to resolve ionizing radiation-induced 53BP1 foci. **A.** Nuclei from irradiated cells 40 min (top)

and 90 min (bottom) post-irradiation. Two cells (red arrows) resolved 53BP1 foci more rapidly and were selected for capture. Scale bar: 14μm. **B, C.** Smoothed normalized histograms showing the fraction of cells detected as a function of the number of induced GFP-53BP1 (B) or endogenous 53BP1 (C) foci for 5 time points. Sorted Populations #1 and #2 resolve foci faster than their parental counterpart as illustrated by the more rapid shift towards the left (zero foci per cell) observed for these two populations. **D.** Illustration of automatic nuclei segmentation and detection of foci (top) and source image (bottom). Objects detected as nuclei are circled in red, segmented foci appear as green circles. Scale bar 25 μm. **E.** Immunoblot showing the amount of 53BP1 at 0, 60, 90 minutes post-irradiation in doxycycline induced cells (+) and non-induced cells (-). 53BP1 levels are not altered in Populations #1 or #2 compared to the parental cells. **F.** Cell cycle profiles of U2OS GFP-53BP1 parental cell lines and two extracted populations. Cultures were induced with Dox for 48h and cell cycle was analyzed by DNA content flow cytometry (see material and methods). Values represent the means +/- SEM of three independent experiments. All focus quantification graphs represent the average of 3 experiments, where in each case at least 200 cells were scored.

iii) The ability to quickly resolve 53BP1 foci is transmitted from parental to daughter cells

We next compared the kinetics 53BP1 focus resolution in Populations #1 and #2 vs. the parental cell population. The resolution of foci was quantified using (i) live-cell imaging of GFP-53BP1 (Figure 17B) and also (ii) following immunostaining with anti-53BP1 antibody (when GFP-53BP1 expression was not induced) to evaluate focus formation involving the endogenous untagged protein (Figure 17C). Images were acquired at 45, 60, 75, 90 and 120 minutes post irradiation with 1Gy for the two populations and the distribution of DNA foci per cell compared with that of the parental cell line. We used MATLAB to program a fully automated algorithm for focus quantification (Figure 17D) and analyzed approximately 1800 cells per time-point. This allowed the unbiased evaluation of large datasets as Figure 17B and C taken together represent the behavior of more than 21,000 cells.

Figure 17B and C shows normalized histograms (probability density functions) of the number of foci per cell at each time-point. Importantly, all three populations exhibited similar numbers of foci per cell 45 minutes after irradiation, indicating that the initial formation of 53BP1 foci is comparable between all cell populations. However, we found that the progeny

of captured cells (Populations #1 and #2) retained the original visually-detected phenotype of fast focus resolution. These cells resolved foci at least 1.5 times more rapidly than parental counterparts, as the median number of GFP-53BP1 foci per cell 60 minutes post-IR for Populations#1 and #2 (17 and 15 respectively) is equal to the median number of foci that parental cells exhibit at 90 minutes post-IR. After 75 minutes these numbers of foci are already statistically different (p-values from student T-tests comparing the parental cells to Populations #1 and #2 are respectively 10^{-75} and 10^{-39}). Such differences in focus resolution dynamics is particularly striking in cells for which the expression of GFP-53BP1 is induced (Figure 17B) but is clearly observable as well using immunofluorescence of the endogenous protein in non-induced fixed cells (Figure 17C).

To rule out the possibility that resolution of 53BP1 foci might be due to increased degradation upon IR or to globally decreased levels of the protein, we monitored 53BP1 levels by immunoblotting at different time points post-IR. No changes in the levels of either endogenous 53BP1 or GFP-tagged version was observed (Figure 17E). Finally, FACS analysis shows that all populations exhibit similar ratios of cells in each cell cycle phase (Figure 17F). Therefore, the observed focus resolution differences between populations is unlikely to be attributable to cell cycle-related effects.

iv) Cells can be purified based on morphology

We next sought to illustrate of the utility of scMOCa to capture cells based on their morphology, which have so far proven challenging to sort using currently available technologies. For example, multinucleated cells constitute a rare subpopulation[167-168] that does not express specific markers and cannot be differentiated from mononucleated polyploid cells using DNA-specific stains in a FACS experiment. However, multinucleated cells can be easily identified visually even without DNA staining. In the context of cancer, such cells have been (i) described as generally being more aggressive and metastatic than mononucleated counterparts, and (ii) proposed to be prone to acquisition of drug resistance and cancer relapse[167, 169-171]. Moreover, even though multinucleated cells do not undergo classical

cytokinesis, they can generate mononucleated progeny by budding[167, 170] and influence neighboring cells by secreting factors that promote stemness, as well as by transmitting sub-genomes[167].

Multinucleated cells were isolated using scMOCa and kept in culture for 4 days to evaluate their viability and metabolic activity (Figure 18). We used WGA-alexa647 to stain plasma membranes, and Hoechst for the nuclei (Figure 18) and Mitotracker green FM to tag polarized mitochondrial membranes, indicating that scMOCa preserves the viability of isolated cells (see supplementary figure 18).

As another example of a visual phenotype that can be sorted using scMOCA, we evaluated the differentiation of 3T3 cells into adipocytes. These cells are amongst the most common models to study metabolic disorders, e.g., obesity[172-173]. When cultured for two days in medium containing dexamethasone, insulin and isobutylmethylxanthin (IBMX), an inhibitor of cyclic nucleotide phosphodiesterases, and 3 days in medium containing insulin, a fraction of 3T3 cells differentiate and lipid vesicles accumulate in their cytoplasm. In order to obtain pure adipocyte cultures, flow cytometry sorting based on granularity requires several steps to select cells of interest and then remove false positives, such as debris and cell aggregates[174], whereas scMOCA may provide a much simpler approach to isolated live adipocytes, especially when these are present in very low abundance. We used scMOCA to capture differentiated adipocytes and then kept them in culture for a week (Figure 18). Sorted cells remained viable and maintained their ability to store lipids in vesicles that appear as clear spheres on figure 18, while the magnetic beads that remained attached to cells membranes appear as dark spheres.

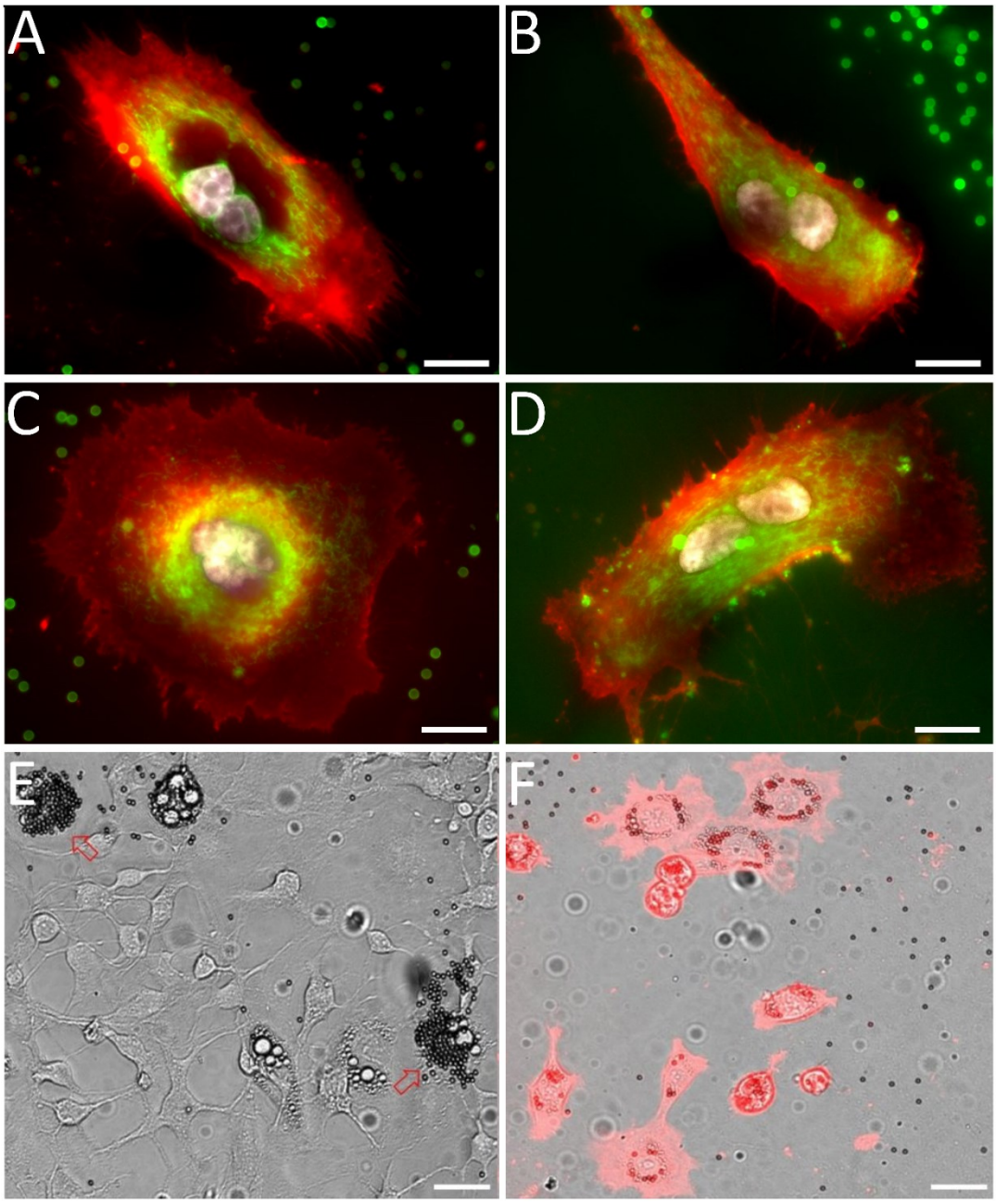


Figure 18. Examples of sorted multinucleated H226 cells one (**A, C**) and two (**B, D**) days after scMOCa. Active mitochondria (Mitotracker) appear in green, plasma membrane (WGA-Alexa 647) in red, an nuclei (Hoechst) in white. Scale bar: 15µm. **E:** 3T3 cell population partially differentiated into adipocytes. Two cells (pointed by arrows) have been tagged with magnetic beads. Three other cells are also differentiated in adipocytes in this field of view but were not selected. **F:** Cells were captured using scMOCA and kept in culture for 6 days before imaging. Cells were stained with WGA-Alexa 647 to highlight membranes. Small black circles are magnetic beads while lipid vesicles appear as small clear circles. Scale bar: 30µm.

e) Discussion

To the best of our knowledge, scMOCa is the only technology that permits isolation, and subsequent clonal expansion, of extremely small numbers of cells from relatively large heterogeneous populations based solely on visual criteria. scMOCa is highly efficient, as the fraction of tagged cells collected in the top chamber exhibits minimal capture losses and high specificity. Rare false positive cells, presumably attached by cell junctions to true positive cells, can be eliminated by repeating the sorting procedure to reach 100% purity. The most widely used cell sorting technique, FACS, is not optimized for sorting rare cells. Adaptations needed for capturing cell populations representing <1% of the sample with high specificity make FACS experiments cumbersome and inefficient. Moreover, repetition of flow cytometry sorting to obtain pure samples of a given cell type imposes can only be performed with robust cell types due to reduced survival and proliferation capacity[17]. More refined procedures have been developed to sort rare cells via binding to microfluidic channels coated with antibodies against specific surface markers of interest[33]. However, this requires high affinity antibodies that are specific to the target cell types and leads to dilution of cells in laminar flows within microfluidics chips[175-176], which can become a drawback for downstream applications. Techniques based on magnetism display an improved capacity to isolate rare cells without dilution[158]. Nevertheless, while the majority of protocols that use magnetic fields can capture cells of interest with efficiency near 90%, their specificity remains a major challenge, as published results vary between 10 and 80% purity for captured cells[177], generally closer to 50%[34, 159, 178-179]. Finally, only a handful of approaches allow label-free cell sorting, where intrinsic physical properties, such as size[180-181] or magnetic susceptibility[34, 175, 182] differentiate the target population. Filtration, for example, relies on porous membranes to capture cells based on size and deformability[183-184] and can achieve 80% efficiency. Dielectrophoresis exploits natural differences in dielectric properties of cell types for discrimination and circulates cells in microfluidics channels, deviating target cells within an electric field[22, 24].

The application we introduced here is focused on magnetic separation, but the same concept of adding particles to individual live cells may open the door to novel strategies where

other actionable properties can be exploited in a simple and straightforward manner. For example, fluorescence or electron density can be manipulated on single cells[46], and recent advances in cellular nanotechnologies such as scattering and plasmon resonance using gold nanoparticles, thermal capacity with nanoshells, or electrical properties using carbon nanotubes can now be modulated only on chosen cells using low-cost commercially available reagents.

ScMOCa presents critical advantages over more traditional sorting techniques. It allows isolation of live cells without previous knowledge of surface markers and can simply be based on morphological traits such as the presence of nuclear foci or lipid vesicles and the number nuclei. More importantly it has the potential to sort based on time-dependent characteristics such as migration speed or foci resolution. In addition, because sorting is carried out in small chambers of similar size, there is no sample dilution. This prevents cells from sustaining strong shear stress upon passing through microfluidic tubing[177], and allows their use in downstream applications such as cell culture, reinjection, or even lysis prior to transcriptomic or proteomic analysis. ScMOCa crosslinks biotin to cell membrane and the strength of the ensuing biotin-streptavidin bond is extremely high (K_d = 10^{-15}M). In comparison, the bonds utilized in immunochemistry are much weaker, from 10^{-12} to 10^{-9} [185-186], which may cause tags to detach from cells because of shear stress within the microfluidics tubing[187]. Another example is provided by ligands targeting the major histocompatibility complex (MHC) on immune cells where binding strength is so weak that ligands usually need to be grouped in tetramers for increased strength[188-189]. Finally, while the precise mechanisms influencing 53BP1 focus resolution was not investigated in our proof-of-principle experiments, our data demonstrates that markers used for identification need not be exposed on the membrane since the spatial distribution of fluorescent signal originating from the nucleus were used here as a reporter.

Simplicity is a key advantage of scMOCa, as it does not require highly specialized software, or hardware such as microfluidic chips. Indeed, a standard confocal microscope with no modification, simple handmade chambers and low-cost magnets are all that is needed to sort single cells of choice from among tens of thousands. The main limitation of scMOCa is that

high throughput implementations would depend on efficient image processing tools for cell detection. While automated detection and tagging are possible on motorized microscopy systems, the duration of the procedure is roughly proportional to the number of target cells. Thus, even if laser illumination of a single cell typically requires one second, this might become a limitation for applications that deal with large cell numbers.

The capacity of scMOCa to isolate and profile individual cells within a large population based purely on visual phenotypes constitutes a powerful tool for understanding cellular heterogeneity. We envision that one potential application of high interest would combine scMOCa with single cell sequencing to characterize the molecular basis of differential metastatic potential among particular cells within a tumour[190-195]. Indeed, scMOCa can easily be combined with currently available techniques that allow sequencing RNA from single cells captured in wells[196] and microfluidic chips[197-198]. More generally, it is becoming increasingly obvious that the capacity to analyze rare cells in heterogeneous populations will be useful in designing personalized treatments for cancer[199-200] as well as for inflammatory, autoimmune, and neurologic disorders[177, 201-202].

f) Materials and methods

i) Key resource table

Reagent type (species) or resource	Designation	Source or reference	Identifiers
cell line (Homo sapiens)	U2OS	ATCC	RRID: CVCL_0042
cell line (Canis familiaris)	MDCK	ATCC	RRID: CVCL_0422
cell line (Mus musculus)	IMCD	ATCC	RRID: CVCL_0429
cell line (Homo sapiens)	h226	ATCC	RRID: CVCL_1544
cell line (Homo sapiens)	LF-1	Dr John Sedivy	RRID: CVCL_C120
cell line (Homo sapiens)	HUVECS	ATCC	TCC® PCS-100-013™
cell line (Mus musculus)	3t3-L1	ATCC	RRID: CVCL_0123
chemical compound, drug	IMBMX	sigma-aldrich	cat #: I5879-100MG
chemical compound, drug	Dexamethasone	sigma-aldrich	cat #: D1756-25MG
chemical compound, drug	magnetic beads	thermofisher	cat #: 65305
chemical compound, drug	b4f	sigma-aldrich	cat #:B9431-5MG
Commercial assay or kit	2X SYBR Green Master Mix	Bimake	cat #: B21203
antibody	rabbit anti-53BP1	Santa-cruz	cat #: sc-22760
antibody	rat anti-tubulin	Abcam	cat #: ab6161

Tableau II. Key resource table

ii) Cell culture

U2OS osteosarcoma cells, MDCK (dog) cells, and IMCD (mouse) cells were grown in DMEM/F12 medium supplemented with 10% FBS and antibiotics, all purchased from Thermofisher Scientific. One day prior to the experiment, cells were detached and seeded on either collagen-coated glass coverslips or circular pieces of Aclar (polychlorotrifluoroethylene) coated with collagen, onto which polydimethylsiloxane (PDMS) chambers had been placed (see below).

A U2OS cell line with inducible expression of GFP-tagged 53BP1 was constructed as previously described[203] using pcDNA5-FRT/TO-eGFP-53BP1[165] (Addgene plasmid #60813) and the U2OS Flip-In TREX host cell line[204] (both generous gifts from Dr. Daniel Durocher, University

of Toronto). Cells were selected in medium supplemented with 200µg/mL hygromycin and 5µg/mL blasticidin. GFP-53BP1 expression was induced by addition of 5 µg/mL doxycycline for 48 hours.

H226 cells were grown in RPMI medium supplemented with 5% FBS and antibiotics (Thermofisher Scientific). Four days prior to the experiment, cells were exposed to 6 µg/mL cytochalasin B for 24 hours. Low-passage primary human lung fibroblasts (LF-1) were a kind gift from Dr John Sedivy[205]. Cells were grown in Eagle's MEM (Corning) containing 15% FBS, essential and nonessential amino acids, vitamins, L-glutamine, and antibiotics (Life Technologies). HUVECS were grown in Endogro TM (Millipore) supplemented with VEGF. Primary dorsal root ganglion (DRG) neurons were harvested from IsI-Gcamp6 x TRPV1-cre mice and cultured in plastic bottom dishes (as detailed elsewhere[206]) one day prior to the sorting.

iii) 3T3-L1 cell culture and adipogenic differentiation

Pre-adipocyte 3T3-L1 cells were grown in DMEM medium supplemented with 10 % FBS (Gibco), 2 mM glutamine (Wisent) and 1 % Penicillin/Streptomycin (Biobasic). For adipogenic differentiation of 3T3L1, the cells were plated at confluency and media was changed to induction media containing 10 % FBS, 1 % Penicillin/Streptomycin, 1 µM Dexamethasone, 1 µg/ml Insulin and 500 µM IBMX (Sigma). Two days post-induction, the medium was changed to maintenance media containing 10 % FBS (Gibco), 1 % Penicillin/Streptomycin (Biobasic), 1 µg/ml Insulin. After three days post-induction, 10 000 cells were plated on homemade chambers for sorting.

iv) Mouse Embryonic Stem cell (mES) culture

mES cells were grown in DMEM medium supplemented with 15 % FBS (embryonic stem cell qualified, Wisent), 1 X non-essential amino acids (Sigma), 100 µM 2-Mercaptoethanol (Gibco), 1000 Units/mL Leukemia inhibitory factor (LIF, Stemcell), 2 mM glutamine (Wisent) and 1 % Penicillin/Streptomycin (Biobasic) on 0,1 % porcine gelatin coated plastic dishes (Sigma). About 10 000 cells were plated for sorting as above.

v) PDMS chambers

PDMS chambers were prepared by pouring a mix of resin and curing agent (10:1 ratio) in a petri dish to achieve a gel thickness of 2 mm. The dish was degassed overnight in a vacuum chamber and the resin allowed to polymerize at room temperature for 2 days. Square pieces were cut with a blade, circular wells of 5 mm diameter were made using a biopsy punch from Miltex® (33-38) (see Figures 1 B and C) and placed on either glass or Aclar coverslips (onto which PDMS naturally adheres).

vi) scMOCa protocol

Cells were incubated in regular medium with 40 µg/mL biotin-4-fluorescein (Sigma) on glass coverslips or Aclar substrates. A spot within each cell of interest was illuminated at 473 nm with the laser of a confocal microscope at 75 µW for 2 seconds with 10X 0.4NA objective. The sample was then thoroughly rinsed in PBS, and medium containing 8 µL of streptavidin-coated ferromagnetic beads of 2.8 µm in diameter (Thermofisher, 65305 and 11533D) was added. When beads were attached to a whole area rather than a single cell (Figure 1B and figure 1-supplement 1) the sample was scanned with a 700 µW laser scanned at 0.2 mm/s with a 0.4 NA objective in a succession of lines 0.005 mm apart to form a pattern generated from a binary image.

Beads were pulled down in contact with the cells and re-suspended 3 times, attracted by a magnet placed alternatively below or above the sample. Cells were then rinsed 3 times with PBS and a magnet was positioned above the sample to remove unbound beads. After this, very few beads remain in the dish (Figure 1C).

Cells are detached using 0.25% trypsin (Thermofisher, 25200072) for magnetic capture. The resulting cell suspension is then subjected to a magnetic field that attracts positive cells upwards to a collection chamber, while negative cells settle by gravity in the original chamber, regardless of the total number of cells in the sample.

More specifically, once the original PDMS culture chamber contains a suspension of individual cells in trypsin, a second identical PDMS culture chamber is placed on top of the first

one as depicted in Figure 2A. The structure that holds the top chamber in place can be built with Lego bricks (Figure 1-supplements 3 and 4): the collection chamber is positioned between two Lego bricks that maintain it at 6 mm above the cells (Figure1C). While magnetic attraction of tagged cells towards the collection chamber is quick, negative cells require four minutes to settle down to the original chamber before the top chamber is separated, flipped, and the magnets removed. This procedure needs to be performed slowly to minimize turbulence and to avoid capture of negative cells.

These manipulations are repeated three times to attain maximum specificity (Figure 2C). The collection chamber is always filled with trypsin solution to avoid rapid cell adhesion, and gentle up and down pipetting can be performed to prevent cell clumping. Only for the last capture is the collection chamber filled with medium in which the cells will be expanded. The entire procedure is summarized in figure 2C.

Experimental conditions need to be fine-tuned for different cell types. The most important parameters that need to be optimised are surface coating of both donor and collection chambers, duration and number of repeats of the sorting steps. The collection chamber should provide optimal plating efficiency to maximize cell survival of very few cells while the donor chamber should allow strong adhesion of the cells to allow thorough rinsing of free magnetic beads. In our experience collagen coating provides strong cell attachment, but also generates extracellular fibers where beads and negative cells can be entangled and captured. Gelatin solves the issue of collagen fibers, but cell adhesion is slightly reduced, which may cause cell loss during rinsing. Uncoated substrates are an easy solution for cells like U2Os but many cell types including primary cells do not proliferate well on such surfaces. Plastic bottom chambers allow better cell adhesion and survival, but their reduced optical quality may hamper the precise observation of selection criteria. In this respect Aclar possesses excellent optical properties and represents an excellent alternative. For most cell types, longer incubations (approximately 4 minutes) allow negative cells to settle down in the donor chamber, reducing the number of repeats required for optimal purity. On the contrary,

experimentation with cells that adhere rapidly (e.g. MDA-MB-231), require the capture protocol to be performed as quickly as possible and more repeats may be needed.

In our hands the best results were obtained using 10 magnets each generating a 1.2 Gauss magnetic field and 2 mm deep PDMS wells. In this condition, it is important that the distance between the bottom of each chamber is kept at 6 mm to allow the magnetic field to attract all tagged cells against gravity to the collection chamber while preventing the turbulence generated by the separation of the chambers to bring negative cells into the collection chamber. Increasing this distance requires a stronger magnetic field, which in turn reduces viability of captured cells. The diameter of the chambers should also be 5-6 mm, to ensure the necessary surface tension that allows merging and splitting the media in both donor and collections chambers.

vii) Cell sorting using commercial magnetic cell separation columns

30 000 U2Os cells were plated in our homemade chambers one day prior to sorting. On the day of the experiment, 30 cells were arbitrarily chosen and tagged in 3 independent experiments. We manually counted and verified that the right number of cells (30) were covered with magnetic beads in each dish. Commercial MACS columns were washed with PBS containing 0.5% BSA and 2 mM EDTA as indicated by the manufacturer. Cells were detached using 60 µL trypsin and then diluted in 500 µL of the same buffer and placed in the column in the magnets from Miltenyi Biotec. Columns were rinsed three time with buffer, then removed from magnets and washed with 5 mL buffer. Cells were then centrifuged, resuspended in 70 µL medium and placed in new homemade chambers for observation and counting under the microscope. Any cell that had visible magnetic beads on its membrane was considered as a positively selected cell, while cells free of beads were counted as negative cells.

viii) Identification and isolation of "fast resolving" live cells

Forty-eight hours after induction of GFP-53BP1 with doxycycline, U2OS cells were irradiated with 0.5 Gy of IR. A first set of images was acquired with a 40X, 0.95 NA objective 45 minutes post irradiation, to detect focus formation.

Cells that displayed a > 85% reduction in the number of foci at the second time point (2 hours) were considered “fast-resolving”. Biotin-4-fluorescein (0.04 mg/mL) was then added to the medium, and such cells were illuminated for 2 seconds through a 10X 0.4NA objective with 75 µW of laser intensity at 473 nm.

ix) Immunofluorescence and automated detection of nuclear GFP-53BP1 foci

Immunofluorescence was performed to evaluate levels of endogenous 53BP1 foci. Briefly, cells were rinsed with PBS, and fixed 15 min with 4% paraformaldehyde in PBS. Cells were then permeabilized for 10 min with 0.5% Triton X-100 in PBS, rinsed twice in PBS and twice in PBS + 0.05% Tween-20 and then blocked in PBS + 3% BSA and 0.05% Tween20. Rabbit anti-53BP1 antibody (Santa-Cruz) was diluted 1:500 and incubated on the cells for 3 hours. Cells were rinsed in PBS + 0.05% Tween-20 and incubated with Alexa-488 anti-rabbit for 1 hour, washed 3 additional times, and finally imaged for focus quantification.

An image processing pipeline was programmed to fully automate DNA focus detection as we have previously done[207-208]. Cell nuclei were detected using the background signal of remaining free GFP-53BP1 protein by Otsu thresholding[209]. This initial detection was used to create a mask, where objects were filtered for their size, signal saturation, and shape. A band-pass filter was used to enhance the signal generated by objects the size of a 53BP1 focus. Local maxima were then detected using a threshold automatically calculated for each nucleus.

x) Mitochondria staining and imaging

Sorted multinucleated H226 cells were stained 2 and 4 days after their isolation. Mitotracker green FM (Thermofisher Scientific, M7514) was used at 150 mM for 20 minutes, followed by a 5 minutes incubation in Hoechst 33342 to stain nuclei, and WGA-alexa 647 to stain plasma membranes. Images were acquired with a 60X 1.35NA objective.

xi) Imaging

Cell selection and CLaP were performed on an Olympus IX71 microscope (Olympus Corp.) with the appropriate epifluorescence filters, in medium at 37 degrees, 5% CO2, with a 10X 0.4NA objective and an Orca Flash 4.0 camera (Hamamatsu Photonics).

Images of irradiated GFP-53BP1 expressing cells were taken at two time points using a 40X, 0.85NA objective and compared to identify outliers. Since laser tagging was performed with a 10X 0.4NA objective, cells were identified in a new live image at different magnification during tagging.

Automatic acquisition of immunostained samples for characterization of large numbers of cells from purified cell populations was performed with an automated Zeiss AxioObserver Z1 Epifluorescence microscope, at room temperature in PBS with Zen Blue software and a 20X 0.85NA objective.

xii) Cell cycle analysis

Exponentially growing cell cultures were trypsinized, fixed with 70% ethanol, and stored at -20°C until use. Fixed cells were washed with PBS and treated with 0.5% triton X-100 for 10 minutes at room temperature. After washing with PBS, cells were resuspended in PBS containing 2 µg/ml propidium iodide and 0.2 mg/mL RNase A and incubated for 30 minutes at room temperature. Samples were analysed by flow cytometry on a FACSCalibur instrument (Becton-Dickinson). Data was analyzed with FlowJo v10 software, and cell cycle phases were determined using the Watson algorithm.

xiii) Conditioned medium

U2OS cells were plated at a density of 2 million cells per 10 cm dish. 24 hours later, medium was removed and filtered through a 0.2 µm filter to ensure sterility and remove any floating cells. Conditioned medium was always prepared fresh.

xiv) Polymerase chain reaction

10 cells were resuspended in 40 µL of water and boiled for 10 minutes. Samples were subjected to 24 PCR cycles using Agilent Herculase II with primer sets specific for the mitochondrial gene *Cytb* of either dog or mouse. 2µL of each reaction were then used for PCR or qPCR with each primer set. Total genomic DNA from either dog or mouse cells were used as controls. The primers used are *Cytb1L*(5'- CATAGCCACAGCATTCATGG -3'), *Cytb1R*(5'- GGATCCGGTTTCGTGTAGAA -3'), and *Cytb2L*(5'- CCTCAAAGCAACGAAGCCTA -3'), *Cytb2R*(5'- TCTTCGATAATTCCTGAGATTGG -3'), which respectively amplify fragments of 247 nt and 196 nt from the *Cytb* gene of dog and mouse. Quantitative PCR was performed with the above primer pairs using the 2X SYBR Green Master Mix (Bimake) and an ABI7500 instrument (ThermoFisher). The amount of dog and mouse DNA in each sample was calculated using standard curves made from serial dilutions of genomic DNA isolated from each cell type.

xv) Immunoblotting

Immunoblotting was performed with total cellular extract using standard protocols. Antibodies used were rabbit anti-53BP1 (Santa-Cruz, sc-22760) and rat anti-tubulin (Abcam, ab6161).

g) Acknowledgments

We thank Maryam Tabatabaei and Sébastien Talbot (Université de Montréal) for help with experiments using DRG cells. This work was supported by grants from the Natural Science and Engineering Research Council of Canada to E.D., E.B.A, S.C and H.W, Genome Canada/Génome Québec and Canadian Cancer Society to SC, Fonds de Recherche du Québec – Nature et Technologies to SC, and the Canadian Institutes of Health Research to E.D., E.B.A. and H.W. S.C., E.B.A. and H.W. hold salary awards from the Fonds de Recherche du Québec – Santé.

h) Competing interests

The authors declare no competing financial interests.

Chapitre 8 : Discussion et conclusions

a) Synthèse des travaux

Dans cette thèse, nous avons dans un premier temps exposé le développement d'une technique de fonctionnalisation de membranes cellulaires assistée par laser qui permet d'attacher des étiquettes moléculaires sous forme de streptavidine colorée afin d'attribuer une identité à des cellules uniques. Cette étiquette permet de reconnaître les cellules aussi bien au sein de l'échantillon lors d'observations subséquentes que sur des plateformes de séquençage telles que C1^{TM}. Nous avons montré que la procédure n'endommage pas les cellules qui restent viables et prolifèrent normalement, sans altération de leur transcriptome. Cette marque cellulaire peut aussi être utilisée dans des échantillons tridimensionnels, y compris ex-vivo soit avec une excitation classique en 1-photon pour étiqueter toute une colonne de tissu, soit avec une excitation 2-photons pour sélectionner une tranche de tissu, ou même pour taguer une seule cellule d'un échantillon tridimensionnel. Plusieurs couleurs peuvent être utilisées simultanément pour être combinées en codes-barres afin d'identifier davantage de cellules. Chaque couleur ajoutée augmente significativement le nombre de tags attribuables puisque celui-ci varie selon la formule $2^{(\text{nombre de couleurs})}$. Nous avons référencé cinq paires de molécules qui pourraient théoriquement être utilisées simultanément pour générer un maximum théorique de 2^5-1=31 codes couleurs et nous avons démontré le principe de ce multiplexage avec trois paires de molécules.

De plus, CLaP nous a permis de coller transitoirement des cellules choisies à leur support de culture, nous donnant l'opportunité de décoller et rincer toutes les autres cellules. Cette utilisation est prometteuse à la fois pour trier des cellules, en ne gardant que les cellules d'intérêt, et pour contrôler avec précision le positionnement de cellules sur leur substrat. Des applications comme l'étude d'interactions entre types cellulaires, le génie tissulaire, ou l'utilisation de cellules nourricières pour la culture de cellules plus sensibles peuvent grandement bénéficier d'une telle capacité.

Enfin, une tierce molécule biotinylée peut être liée à la membrane fonctionnalisée afin d'être délivrée de manière ciblée, ce qui pourrait être utilisé pour modifier génétiquement des cellules choisies ou leur délivrer des molécules effectrices telles que des siRNA, des virus, des liposomes…

Cette nouvelle méthode trouve des applications dans des situations très variées dans lesquelles les anticorps ne permettent pas d'identifier en fluorescence les cellules d'intérêt parce que le critère de sélection n'est pas biochimique. Il n'y a ainsi pas de marqueur spécifique de la position ou la vitesse de déplacement de la cellule, de ses contacts avec son environnement, de sa taille ou encore de son nombre d'organelles. De plus, il arrive que le signal utilisé pour identifier les cellules d'intérêt résulte d'une trans-localisation d'un signal fluorescent sans changement de l'intensité totale du signal comme dans notre exemple de la résolution des foci de la protéine 53bp1. Dans de tels cas de figure, le microscope permet de reconnaître aisément les cellules cibles, mais les plateformes de tri comme FACS ne lisent que l'intensité totale, et ne sont pas sensibles à la position du signal dans la cellule. De même, si le signal est intracellulaire, il n'est pas atteignable par tous les types de sondes moléculaires. Une simple observation au microscope permet d'observer chacun des critères cités ici, et de les utiliser pour choisir les cellules à étiqueter avec CLaP. Nous avons montré dans l'article *« Live single cell laser tag* » que les cellules taguées peuvent être triées en utilisant un FACS. Cependant, la nature de la méthode, au cours de laquelle les cellules sont étiquetées une à une, fait que les cellules à trier sont intrinsèquement rares, indépendamment de la proportion des cellules cibles dans l'échantillon de départ. Nous avons cependant montré que CLaP peut être utilisé pour identifier des cellules sur les plateformes de séquençage d'ARN de cellules uniques comme C1. Néanmoins, ces techniques requièrent de capturer individuellement les cellules. Cette étape est très peu efficace puisque seulement 10% des cellules sont effectivement capturées. Or les cellules qui nous intéressent étant très rares, elles sont précieuses et une technique de tri idéale doit permettre leur isolation sans pertes. Le tri de cellules rares est encore aujourd'hui un défi. Les techniques classiques comme FACS ou MACS donnent des résultats peu satisfaisants en termes de pureté et de nombre de cellules effectivement triées.

Dans la deuxième partie de cette thèse, nous avons présenté comment CLaP n'est pas limité à la simple identification de quelques cellules parmi des milliers mais peut aussi servir à les isoler. Nous montrons qu'il est possible de fonctionnaliser une à une quelques cellules pour lier à leur membrane des billes magnétiques. Des cellules viables et pures sont alors facilement extraites en utilisant un simple aimant. Des populations de cellules fragiles, comme des HUVECs, des fibroblastes, des neurones primaires ou des cellules souches embryonnaires peuvent être triées sans dommages. Nous avons illustré l'efficacité de la technique en générant des sous populations de cellules sélectionnées pour leur vitesse de résolution des foci de la protéine 53bp1 un marqueur associé aux cassures double brin induites sur l'ADN. Nous avons choisi d'être très sélectifs et de ne garder que les rares cellules résolvant le plus rapidement les foci. Nous avons maintenu ces cellules en culture pour générer de nouvelles lignées qui ont conservé leur capacité à résoudre rapidement les foci. D'autre part, nous avons isolé des cellules multinucléées. Ces cellules apparaissent spontanément dans les cultures de cellules et dans les organismes[210-212]. Elles jouent un rôle déterminant dans le développement de la résistance aux médicaments. Dans le cas du cancer notamment, elles peuvent survivre aux traitements, et causer des rechutes à la fin du traitement en se divisant de manière asymétrique, générant ainsi de nouvelles cellules de cancer[167]. Les cellules que nous avons isolées restent viables en culture et peuvent être utilisées pour une plus ample caractérisation de leur biochimie. Ainsi, des cellules rares viables peuvent être identifiées par observation au microscope, être couvertes de billes magnétiques, et isolées. Alors que leur capture reste un défi, les cellules rares sont le centre d'intérêt de domaines aussi variés que le développement des cancers et des métastases, les diagnostics prénataux, l'isolation de cellules souches rares, de cellules immunitaires ou infectées et enfin dans les études des dynamiques de population et notamment de l'apparition des résistances aux médicaments.

Enfin, une application particulièrement encourageante de CLaP est son utilisation pour reconnaître des cellules lors du séquençage de cellules individuelles. Des progrès très récents permettent de connaître la séquence des bases constituant les molécules d'ARN messagers à l'intérieur de cellules individuelles. Cependant, ces nouvelles techniques requièrent généralement de digérer l'échantillon en une suspension de cellules individuelles qui sont

ensuite capturées soit dans des puits soit dans des gouttes. La destruction de l'intégrité du tissu cause une perte totale de l'information spatiale : la position de chaque cellule capturée au sein de l'échantillon est perdue, et les potentielles interactions entre cellules ne peuvent donc pas être étudiées. La technologie C1^{TM}, qui permet une observation au microscope des cellules capturées sur une puce de micro fluidique, est utilisable pour reconnaître un code couleur qui aurait été au préalable attribué par CLaP afin d'encoder la position des cellules dans l'échantillon. Cependant le taux de capture est faible, et le nombre total de cellules capturées et séquencées est limité. L'autre approche, qui semble devenir le standard, consiste à créer une suspension de cellules qui sont encapsulées dans des goutes à l'intérieur desquelles elles sont lysées, et leur ARNm est capturé par des amorces encodées. Cette technique ne permettant pas la visualisation des cellules, CLaP ne peut être directement utilisé pour attribuer un code couleur à la position des cellules. Une autre technique, publiée en 2016, propose de lyser les cellules sur une puce contenant des amorces encodées[121]. Ces amorces permettent de capturer les ARNms comme le font les billes utilisées dans Dropseq. L'origine spatiale de chaque molécule de la librairie générée sur ces puces est donc encodée dans la séquence de l'ADNc. La fabrication de la puce et la diffusion des molécules ne permettent cependant pas d'obtenir une résolution spatiale de l'ordre de la cellule.

Considérant l'augmentation très rapide du nombre d'articles utilisant ces techniques de séquençage, il nous a paru nécessaire de développer une méthode permettant d'encoder l'identité de chaque cellule dans son environnement. Nous avons modifié CLaP pour attribuer aux cellules un identifiant sous la forme d'une chaîne de nucléotides qui peut être lue en même temps que le transcriptome sur les plateformes comme Drop-seq. Cet oligonucléotide est fixé avec un laser à la membrane de cellules choisies. Il n'est pas décroché lorsque l'échantillon est digéré en une suspension de cellules et est libéré plus tard, lors de la lyse cellulaire. Ce tag permet d'identifier les transcriptomes de cellules d'intérêt tout en bénéficiant de la profondeur de lecture des récentes technologies de séquençage de CI. La capacité de marquer une cellule arbitraire afin de la retrouver après le séquençage du tissu, malgré l'homogénéisation requise par les protocoles de séquençage de cellules individuelles est une grande avancée technologique permettant de lier l'information transcriptionnelle à la

localisation de chaque cellule dans l'organisme étudié. En effet, s'il existe aujourd'hui de plus en plus de méthodes permettant le séquençage de cellules individuelles, la spatialisation des données ainsi générées reste un grand défi. Seuls quelques travaux, recensés dans cette thèse, ont été publiés pour proposer une solution à ce besoin.

b) Travaux futurs

Tous les protocoles d'étude des cellules individuelles (CI) ne requièrent pas l'isolation des cellules d'intérêt. En effet, certains s'appliquent à des suspensions de cellules uniques obtenues par digestion de l'échantillon. Nous avons utilisé l'une de ces techniques, Fluidigm C1™, pour séquencer les cellules taguées par CLaP et les résultats de cette manipulation sont dans l'article présenté au chapitre 3. Cette technologie repose sur l'utilisation de plaques microfluidiques contenant de 96 à 800 sites de capture. Lorsque des cellules en suspension sont injectées dans le système, elles circulent dans des micro-canaux dont la forme est étudiée pour créer une dépression à chacun des sites de capture (Figure 4). Dès qu'un site est occupé par une cellule, la dépression disparait et le reste des cellules passe simplement jusqu'au site suivant. Cette technique présente l'avantage de fournir une très bonne couverture des ARNs présents dans une cellule et la microfluidique permet de limiter les volumes de réactifs utilisés. De plus cette approche comprenant une étape d'observation au microscope des cellules capturées, elle est idéalement adaptée à la reconnaissance des cellules identifiées par CLaP. Cependant, le taux de capture très faible, d'environ 10%, cause la perte d'une majorité des cellules, et le nombre de cellules qui peuvent être analysées est finalement très faible. Les cellules étiquetées individuellement par CLaP représentent nécessairement une très faible proportion (de l'ordre de 0.1%) du nombre total de cellules d'un échantillon à cause du faible débit de la technique. Ainsi, une plaque de 96 (ou même 800) puits ne peut capturer un nombre suffisant de cellules taguées pour obtenir un résultat statistiquement significatif.

D'autres technologies (Drop-seq[35], CITE-seq[48], Abseq[3]) présentées dans la revue au chapitre 2 de cette thèse reposent sur l'encapsulation d'un très grand nombre de cellules dans

des gouttes à l'intérieur desquelles le matériel génétique est extrait et préparé pour le séquençage. Ces méthodes n'étant pas limitées par le nombre de sites de capture présents sur une plaque, elles permettent l'analyse d'un nombre bien plus conséquent de cellules et accroissent les probabilités d'isoler et séquencer une des cellules identifiées par CLaP. Cependant, ces techniques ne permettent pas de conserver l'identité des cellules : il n'est pas possible de reconnaître dans la liste des séquences détectées celles qui correspondent par exemple à la cellule qui migrait plus vite que les autres, à celle qui était en contact avec un agent pathogène visible au microscope, ou encore à celle qui avait une forme allongée. De plus, on ne peut visualiser les cellules capturées, et l'étiquette fluorescente apposée par CLaP sur les membranes des cellules d'intérêt ne peut être lue par ces approches.

Nous proposons d'utiliser CLaP pour accrocher à la membrane plasmique une séquence d'ADN qui servira de code barre pour identifier la cellule. En effet, puisqu'il s'agit d'ADN, cette séquence connue pourra être retrouvée dans la liste des molécules séquencées, et permettra l'identification des cellules taguées.

Afin de remplir leur rôle d'étiquette cellulaire, les molécules d'ADN doivent remplir un certain cahier des charges. En premier lieux, ces codes-barres d'ADN doivent être biotinylées afin d'être attachés à des cellules individuelles en utilisant CLAP : une biotine est liée à la membrane par photoblanchiment, puis une streptavidine permet de créer un lien entre la cellule et l'oligonucléotide biotinylé. D'autre part, ils doivent posséder une queue polyadénylée afin d'imiter la structure des ARNs messagers. En effet, les techniques de séquençage de cellules uniques reposent sur l'utilisation de billes couvertes d'amorces qui capturent et permettent l'amplification des ARN messagers de la cellule. Les amorces immobilisées sur les billes contiennent une séquence poly-T qui capturent la queue poly-A des ARN messagers. Dès lors, l'inclusion d'une séquence poly-A dans nos étiquettes d'ADN permet leur capture par les billes commerciales. L'autre extrémité de l'étiquette doit contenir une séquence connue pour hybrider les amorces nécessaires aux multiples PCR impliquées dans la préparation des librairies. Enfin, nous avons inclus entre ces deux séquences un code barre et un identifiant moléculaire unique. Le premier est l'identité qui est attribuée à chaque cellule

ciblée par le laser. Le second est une séquence aléatoire spécifique de chaque molécule, qui permet de compter le nombre de fragments d'ADN capturés avec la cellule, supprimant ainsi le bruit généré par les multiples amplifications requises dans la préparation de l'échantillon pour le séquençage. Le fragment ADN est apposé sur la membrane plasmique, à l'extérieur de la cellule, et doit en être libéré lors de la lyse cellulaire pour être capturé par les billes. Nous avons pour cela inclus un pont disulfure entre la séquence ADN et la biotine. Les ponts disulfures sont clivés par la solution, réductrice, qui permet la lyse des cellules, et nos codes-barres sont ainsi libérés de la membrane plasmique. Nous notons ici que pendant que nous développions cette technique, des anticorps étiquetés avec de courtes séquences ADN sont apparus sur le marché, dont un anticorps anti-biotin qui pourrait être utilisé dans ce protocole.

Nous avons attaché ces codes-barres à des cellules en culture. Afin de visualiser leur présence, nous avons incubé l'échantillon avec une séquence d'ADN complémentaire du code barre, taguée avec un Alexa fluor-647 (Figure 19). Nous avons imité les conditions réductrices de la solution de lyse cellulaire en incubant l'échantillon dans du dithiothreitol (DTT) ce qui a permis de vérifier que les brins d'ADN se séparent ainsi de la membrane plasmique. (Figure 19) Cette technique a été l'objet d'un brevet temporaire et est actuellement le sujet d'une collaboration avec le laboratoire du Dr Claudia Kleinman.

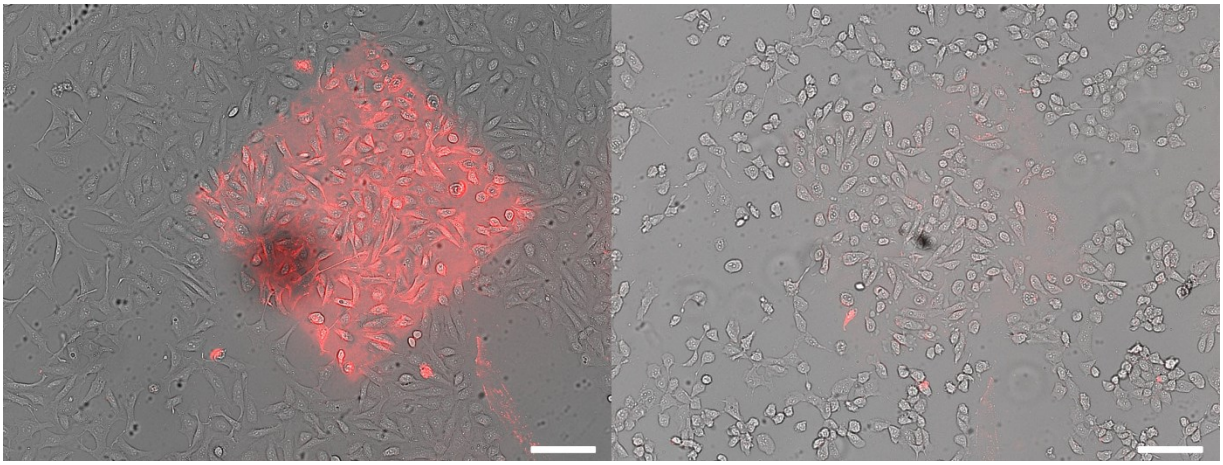


Figure 19. Étiquetage de cellules avec des fragments d'ADN. A Gauche : cellules marquées avec un code barre ADN. Une amorce fluorescente complémentaire du code barre a été utilisée pour montrer sa présence. A droite : le même échantillon après incubation dans du DTT : le code barre est clivé.

c) Limites et améliorations

Les méthodes introduites dans ce travail proposent des outils satisfaisants pour répondre à certaines questions qu'aucune autre approche ne permet d'aborder. Cependant, il est intéressant de noter les limites des techniques présentées ici.

Tout d'abord, nos méthodes n'ont pas encore permis d'identifier et de séquencer des cellules à partir d'un échantillon tridimensionnel. Un certain nombre d'obstacles restent à franchir. La profondeur maximale à laquelle une cellule peut être taguée en deux photons, de l'ordre du millimètre, ne permet pas réellement d'atteindre n'importe quelle cellule dans un être humain. L'utilisation tridimensionnelle reste ainsi plus envisageable dans des modèles de petits animaux tels les poissons, têtards, vers, et souris. Le rinçage et la digestion d'échantillons tridimensionnels accroissent la complexité de la méthode d'une part en augmentant les durées de chaque étape, d'autre part en générant des impuretés dans l'échantillon. Souvent sous forme de fragments de cellule ou de fibres de matrice extracellulaires, ces impuretés sont malheureusement fréquemment non spécifiquement fluorescentes. Un excellent ratio signal sur bruit doit être obtenu pour envisager une utilisation efficace en trois dimensions.

De même nous n'avons pas testé l'utilisation simultanée de plusieurs étiquettes de couleur dans une étude réelle. La qualité du signal généré doit encore être améliorée, notamment en réduisant l'internalisation non spécifique des petites molécules utilisées. L'utilisation de CLaP pour reconnaître l'origine spatiale des cellules lors d'analyses biochimiques peut facilement nécessiter un grand nombre d'étiquettes. En effet, le fait que le nombre d'étiquettes présentes sur une cellule dépende de la surface totale de sa membrane empêche l'utilisation efficace du nombre d'étiquettes détectées comme une identité. Aussi, un encodage des identités des cellules doit être binaire (i.e. présence ou absence du code) et la création d'une hypothétique carte de l'échantillon sur laquelle pourraient être placées toutes les données obtenues par séquençage nécessiterait bien plus que trois étiquettes.

Il est relativement aisé de trouver des partenaires interagissant fortement et spécifiquement entre eux pour générer des marqueurs, d'autant plus que l'on peut utiliser des

fragments d'ADN pour générer un nombre conséquent d'étiquettes ($4^{\text{nombre de bases}}$). Cependant, chacune de ces identités moléculaires doit être associée à un fluorophore que l'on puisse photoblanchir indépendamment des autres. C'est ici que se trouve le principal problème : les fluorophores peuvent en fait être excités, et blanchis, par toute une plage de longueur d'ondes entourant la longueur d'onde d'excitation idéale. Ainsi, leurs pics d'excitation doivent être suffisamment espacés sur le spectre lumineux. Cependant, une seule longueur d'onde pourrait en théorie être suffisante si les étiquettes étaient placées séquentiellement sur les cellules, en remplaçant la solution dans laquelle l'échantillon est immergé entre chaque marquage, comme cela a été fait pour obtenir la Figure 8. Une telle manipulation est néanmoins laborieuse et ne peux pas être envisagée en trois dimensions, les différents temps d'incubation devenant trop longs. Enfin, photoblanchir sélectivement un seul fluorophore dans un échantillon tridimensionnel en utilisant une excitation 2-photons est difficilement réalisable. En effet, l'excitation 2-photons est particulièrement efficace pour photoblanchir un fluorophore, et la plage de longueur d'onde utile est bien trop large pour agir sélectivement sur plusieurs fluorophores.

Enfin, la barrière majeure à franchir pour une utilisation facile à plus grande échelle de la méthode est l'automatisation de la détection et du marquage des cellules. En effet, la procédure dans sa forme actuelle devient rapidement laborieuse dès que le nombre de cellules à étiqueter augmente. Les étapes d'analyse d'image peuvent être automatisées afin de reconnaître les cellules d'intérêt et de contrôler les mouvements du support de l'échantillon et du laser pour assurer leur fonctionnalisation. Un exemple d'analyse automatique a été abordé dans ce travail lors de l'analyse de milliers d'images nécessaire à la caractérisation des populations réparant rapidement les dommages à l'ADN. La difficulté majeure rencontrée, avant même l'analyse individuelle des cellules pour choisir lesquelles doivent être taguées, est leur segmentation. En effet, dans le cadre d'une utilisation automatique, le système doit être capable détecter chaque cellule dans une image de l'échantillon avant d'en faire l'analyse. Pour s'appliquer à notre méthode, cette détection devrait idéalement se faire sans utiliser de marquage fluorescent. Cela permettrait de limiter la toxicité générée par le blanchiment de ce marquage (membranaire), et de conserver un maximum de canaux fluorescents libres pour

l'attribution d'étiquettes (un canal par étiquette) et l'observation de paramètres sur les cellules segmentées (comme par exemple le nombre de foci de 53BP1 dans chaque noyau).

La segmentation d'images de microscopie en lumière blanche est encore aujourd'hui un défi. Elle est particulièrement problématique dans des échantillons dans lesquels les cellules sont confluentes. En effet, beaucoup d'approches permettent de détecter de gros amas cellulaires, mais échouent à détecter la limite entre deux cellules qui se touchent[213-214]. La plupart des approches donnant des résultats probants sont en fait basées sur l'utilisation d'un marquage nucléaire[10, 215]. La connaissance de la position des noyaux permet d'établir le nombre et la position des cellules. Leur membrane peut alors être détectée en faisant croître les objets dans toutes les directions jusqu'à rencontrer la membrane. Cependant, les marqueurs nucléaires interagissent avec l'ADN des cellules et affectent leur viabilité, et leur utilisation condamne l'un des canaux de fluorescence qui ne peut alors pas être associé à une étiquette. Une fois résolu le problème de la détection automatique des cellules à marquer, la principale difficulté technique restante est le placement automatique d'une cellule dans le plan focal de l'objectif. Nous avons testé avec des résultats mitigés des modules commerciaux récents qui permettent de détecter efficacement le plan focal en analysant la réflexion d'un rayon laser sur la surface de culture. Néanmoins, la question de la focalisation automatique dans un échantillon tridimensionnel reste ouverte.

Finalement, l'un des principaux défis lorsque l'on développe une nouvelle méthode de travail est de faire en sorte qu'elle soit acceptée, utilisée, par la communauté. Quel que soit leur impact, ces nouvelles méthodes doivent pour cela être relativement simples à implémenter. Si, dans son principe, CLaP est une approche très simple, chaque nouvelle situation requiert d'ajuster un certain nombre de paramètres. Changer de type cellulaire, de substrat de culture, de type d'étiquette, et même de microscope nécessite de déterminer à nouveau les meilleurs paramètres. Le temps d'illumination, la puissance lumineuse et la concentration en molécules utilisées, la vitesse à laquelle la région d'intérêt est parcourue, ainsi que le traitement des surfaces de culture doivent être optimisés pour chaque problème. Le nombre de paramètres impliqués et l'interdépendance de certains d'entre eux (puissance,

durée d'illumination, concentration en molécules, vitesse de balayage) peut rendre difficile la détermination des paramètres de travail idéaux et décourager un utilisateur appliquant la méthode pour la première fois.

Malgré ces limitations, les méthodes introduites ici sont parmi les premières qui permettent l'étude d'un échantillon directement à partir du phénotype des cellules sans dépendre de leur biochimie. Pour cette raison, je crois qu'elles apportent un jour nouveau sur la façon d'aborder un problème biologique et méritent d'être approfondies.

Références

1. Ibrahim, S. F.; van den Engh, G., Flow cytometry and cell sorting. *Adv Biochem Eng Biotechnol* **2007,** *106*, 19-39.
2. Spitzer, M. H.; Nolan, G. P., Mass Cytometry: Single Cells, Many Features. *Cell* **2016,** *165* (4), 780-91.
3. Shahi, P.; Kim, S. C.; Haliburton, J. R.; Gartner, Z. J.; Abate, A. R., Abseq: Ultrahigh-throughput single cell protein profiling with droplet microfluidic barcoding. *Sci Rep* **2017,** *7*, 44447.
4. Stoeckius, M.; Hafemeister, C.; Stephenson, W.; Houck-Loomis, B.; Chattopadhyay, P. K.; Swerdlow, H.; Satija, R.; Smibert, P., Simultaneous epitope and transcriptome measurement in single cells. *Nat Methods* **2017,** *14* (9), 865-868.
5. Diks, S. H.; Peppelenbosch, M. P., Single cell proteomics for personalised medicine. *Trends Mol Med* **2004,** *10* (12), 574-7.
6. Adan, A.; Alizada, G.; Kiraz, Y.; Baran, Y.; Nalbant, A., Flow cytometry: basic principles and applications. *Crit Rev Biotechnol* **2017,** *37* (2), 163-176.
7. Hogan, S. A.; Levesque, M. P.; Cheng, P. F., Melanoma Immunotherapy: Next-Generation Biomarkers. *Front Oncol* **2018,** *8*, 178.
8. Bendall, S. C.; Simonds, E. F.; Qiu, P.; Amir el, A. D.; Krutzik, P. O.; Finck, R.; Bruggner, R. V.; Melamed, R.; Trejo, A.; Ornatsky, O. I.; Balderas, R. S.; Plevritis, S. K.; Sachs, K.; Pe'er, D.; Tanner, S. D.; Nolan, G. P., Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum. *Science* **2011,** *332* (6030), 687-96.
9. Garcia, M. A.; Nelson, W. J.; Chavez, N., Cell-Cell Junctions Organize Structural and Signaling Networks. *Cold Spring Harb Perspect Biol* **2018,** *10* (4).
10. Goltsev, Y.; Samusik, N.; Kennedy-Darling, J.; Bhate, S.; Hale, M.; Vazquez, G.; Black, S.; Nolan, G. P., Deep Profiling of Mouse Splenic Architecture with CODEX Multiplexed Imaging. *Cell* **2018,** *174* (4), 968-981 e15.
11. Kurosaka, S.; Kashina, A., Cell biology of embryonic migration. *Birth Defects Res C Embryo Today* **2008,** *84* (2), 102-22.
12. Prasad, A.; Alizadeh, E., Cell Form and Function: Interpreting and Controlling the Shape of Adherent Cells. *Trends Biotechnol* **2018**.
13. Villani, A. C.; Satija, R.; Reynolds, G.; Sarkizova, S.; Shekhar, K.; Fletcher, J.; Griesbeck, M.; Butler, A.; Zheng, S.; Lazo, S.; Jardine, L.; Dixon, D.; Stephenson, E.; Nilsson, E.; Grundberg, I.; McDonald, D.; Filby, A.; Li, W.; De Jager, P. L.; Rozenblatt-Rosen, O.; Lane, A. A.; Haniffa, M.; Regev, A.; Hacohen, N., Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. *Science* **2017,** *356* (6335).
14. Zeisel, A.; Munoz-Manchado, A. B.; Codeluppi, S.; Lonnerberg, P.; La Manno, G.; Jureus, A.; Marques, S.; Munguba, H.; He, L.; Betsholtz, C.; Rolny, C.; Castelo-Branco, G.; Hjerling-Leffler, J.; Linnarsson, S., Brain structure. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. *Science* **2015,** *347* (6226), 1138-42.
15. Nguyen, A.; Khoo, W. H.; Moran, I.; Croucher, P. I.; Phan, T. G., Single Cell RNA Sequencing of Rare Immune Cell Populations. *Front Immunol* **2018,** *9*, 1553.

16. Gut, G.; Herrmann, M. D.; Pelkmans, L., Multiplexed protein maps link subcellular organization to cellular states. *Science* **2018,** *361* (6401).
17. Rare-Event Detection and Sorting of Rare Cells. In *Emerging Tools for Single-Cell Analysis*.
18. Navin, N. E., Cancer genomics: one cell at a time. *Genome Biol* **2014,** *15* (8), 452.
19. Seronie-Vivien, S., [Circulating tumor cells: a new challenge for laboratory medicine]. *Ann Biol Clin (Paris)* **2014,** *72* (2), 153-77.
20. Lintz, M.; Munoz, A.; Reinhart-King, C. A., The Mechanics of Single Cell and Collective Migration of Tumor Cells. *J Biomech Eng* **2017,** *139* (2).
21. Baslan, T.; Hicks, J., Unravelling biology and shifting paradigms in cancer with single-cell sequencing. *Nat Rev Cancer* **2017,** *17* (9), 557-569.
22. Gascoyne, P. R.; Noshari, J.; Anderson, T. J.; Becker, F. F., Isolation of rare cells from cell mixtures by dielectrophoresis. *Electrophoresis* **2009,** *30* (8), 1388-98.
23. Issadore, D.; Chung, J.; Shao, H.; Liong, M.; Ghazani, A. A.; Castro, C. M.; Weissleder, R.; Lee, H., Ultrasensitive clinical enumeration of rare cells ex vivo using a micro-hall detector. *Sci Transl Med* **2012,** *4* (141), 141ra92.
24. Hu, X.; Bessette, P. H.; Qian, J.; Meinhart, C. D.; Daugherty, P. S.; Soh, H. T., Marker-specific sorting of rare cells using dielectrophoresis. *Proc Natl Acad Sci U S A* **2005,** *102* (44), 15757-61.
25. Shizuru, J. A.; Negrin, R. S.; Weissman, I. L., Hematopoietic stem and progenitor cells: clinical and preclinical regeneration of the hematolymphoid system. *Annu Rev Med* **2005,** *56*, 509-38.
26. Ye, Q.; Song, D. G.; Poussin, M.; Yamamoto, T.; Best, A.; Li, C.; Coukos, G.; Powell, D. J., Jr., CD137 accurately identifies and enriches for naturally occurring tumor-reactive T cells in tumor. *Clin Cancer Res* **2014,** *20* (1), 44-55.
27. Broz, Miranda L.; Binnewies, M.; Boldajipour, B.; Nelson, Amanda E.; Pollack, Joshua L.; Erle, David J.; Barczak, A.; Rosenblum, Michael D.; Daud, A.; Barber, Diane L.; Amigorena, S.; van't Veer, Laura J.; Sperling, Anne I.; Wolf, Denise M.; Krummel, Matthew F., Dissecting the Tumor Myeloid Compartment Reveals Rare Activating Antigen-Presenting Cells Critical for T Cell Immunity. *Cancer Cell* **2014,** *26* (5), 638-652.
28. Goldberg, J. D., Fetal cells in maternal circulation: progress in analysis of a rare event. *Am J Hum Genet* **1997,** *61* (4), 806-9.
29. Geifman-Holtzman, O.; Makhlouf, F.; Kaufman, L.; Gonchoroff, N. J.; Holtzman, E. J., The clinical utility of fetal cell sorting to determine prenatally fetal E/e or e/e Rh genotype from peripheral maternal blood. *Am J Obstet Gynecol* **2000,** *183* (2), 462-8.
30. Iverson, G. M.; Bianchi, D. W.; Cann, H. M.; Herzenberg, L. A., Detection and isolation of fetal cells from maternal blood using the flourescence-activated cell sorter (FACS). *Prenat Diagn* **1981,** *1* (1), 61-73.
31. Lee, J. S.; Magbanua, M. J. M.; Park, J. W., Circulating tumor cells in breast cancer: applications in personalized medicine. *Breast Cancer Res Treat* **2016,** *160* (3), 411-424.
32. Georgiou, G.; Stathopoulos, C.; Daugherty, P. S.; Nayak, A. R.; Iverson, B. L.; Curtiss, R., 3rd, Display of heterologous proteins on the surface of microorganisms: from the screening of combinatorial libraries to live recombinant vaccines. *Nat Biotechnol* **1997,** *15* (1), 29-34.
33. Nagrath, S.; Sequist, L. V.; Maheswaran, S.; Bell, D. W.; Irimia, D.; Ulkus, L.; Smith, M. R.; Kwak, E. L.; Digumarthy, S.; Muzikansky, A.; Ryan, P.; Balis, U. J.; Tompkins, R. G.;

Haber, D. A.; Toner, M., Isolation of rare circulating tumour cells in cancer patients by microchip technology. *Nature* **2007,** *450* (7173), 1235-9.
34. Zborowski, M.; Chalmers, J. J., Rare cell separation and analysis by magnetic sorting. *Anal Chem* **2011,** *83* (21), 8050-6.
35. Macosko, E. Z.; Basu, A.; Satija, R.; Nemesh, J.; Shekhar, K.; Goldman, M.; Tirosh, I.; Bialas, A. R.; Kamitaki, N.; Martersteck, E. M.; Trombetta, J. J.; Weitz, D. A.; Sanes, J. R.; Shalek, A. K.; Regev, A.; McCarroll, S. A., Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell* **2015,** *161* (5), 1202-1214.
36. Smith, A. M.; Heisler, L. E.; St Onge, R. P.; Farias-Hesson, E.; Wallace, I. M.; Bodeau, J.; Harris, A. N.; Perry, K. M.; Giaever, G.; Pourmand, N.; Nislow, C., Highly-multiplexed barcode sequencing: an efficient method for parallel analysis of pooled samples. *Nucleic Acids Res* **2010,** *38* (13), e142.
37. Tu, J.; Ge, Q.; Wang, S.; Wang, L.; Sun, B.; Yang, Q.; Bai, Y.; Lu, Z., Pair-barcode high-throughput sequencing for large-scale multiplexed sample analysis. *BMC Genomics* **2012,** *13*, 43.
38. Avital, G.; Hashimshony, T.; Yanai, I., Seeing is believing: new methods for in situ single-cell transcriptomics. *Genome Biol* **2014,** *15* (3), 110.
39. Porter, S. N.; Baker, L. C.; Mittelman, D.; Porteus, M. H., Lentiviral and targeted cellular barcoding reveals ongoing clonal dynamics of cell lines in vitro and in vivo. *Genome Biol* **2014,** *15* (5), R75.
40. Lubeck, E.; Cai, L., Single-cell systems biology by super-resolution imaging and combinatorial labeling. *Nat Methods* **2012,** *9* (7), 743-8.
41. Lee, H.; Kim, J.; Kim, H.; Kim, J.; Kwon, S., Colour-barcoded magnetic microparticles for multiplexed bioassays. *Nat Mater* **2010,** *9* (9), 745-9.
42. Li, Y.; Cu, Y. T.; Luo, D., Multiplexed detection of pathogen DNA with DNA-based fluorescence nanobarcodes. *Nat Biotechnol* **2005,** *23* (7), 885-9.
43. Nguyen, H. Q.; Baxter, B. C.; Brower, K.; Diaz-Botia, C. A.; DeRisi, J. L.; Fordyce, P. M.; Thorn, K. S., Programmable Microfluidic Synthesis of Over One Thousand Uniquely Identifiable Spectral Codes. *Adv Opt Mater* **2017,** *5* (3).
44. Braeckmans, K.; De Smedt, S. C.; Leblans, M.; Pauwels, R.; Demeester, J., Encoding microcarriers: present and future technologies. *Nat Rev Drug Discov* **2002,** *1* (6), 447-56.
45. Lovatt, D.; Ruble, B. K.; Lee, J.; Dueck, H.; Kim, T. K.; Fisher, S.; Francis, C.; Spaethling, J. M.; Wolf, J. A.; Grady, M. S.; Ulyanova, A. V.; Yeldell, S. B.; Griepenburg, J. C.; Buckley, P. T.; Kim, J.; Sul, J. Y.; Dmochowski, I. J.; Eberwine, J., Transcriptome in vivo analysis (TIVA) of spatially defined single cells in live tissue. *Nat Methods* **2014,** *11* (2), 190-6.
46. Binan, L.; Mazzaferri, J.; Choquet, K.; Lorenzo, L. E.; Wang, Y. C.; Affar el, B.; De Koninck, Y.; Ragoussis, J.; Kleinman, C. L.; Costantino, S., Live single-cell laser tag. *Nat Commun* **2016,** *7*, 11636.
47. Zheng, G. X.; Terry, J. M.; Belgrader, P.; Ryvkin, P.; Bent, Z. W.; Wilson, R.; Ziraldo, S. B.; Wheeler, T. D.; McDermott, G. P.; Zhu, J.; Gregory, M. T.; Shuga, J.; Montesclaros, L.; Underwood, J. G.; Masquelier, D. A.; Nishimura, S. Y.; Schnall-Levin, M.; Wyatt, P. W.; Hindson, C. M.; Bharadwaj, R.; Wong, A.; Ness, K. D.; Beppu, L. W.; Deeg, H. J.; McFarland, C.; Loeb, K. R.; Valente, W. J.; Ericson, N. G.; Stevens, E. A.; Radich, J. P.; Mikkelsen, T. S.; Hindson, B. J.; Bielas, J. H., Massively parallel digital transcriptional profiling of single cells. *Nat Commun* **2017,** *8*, 14049.

48. Stoeckius, M.; Smibert, P., CITE-seq. **2017**.
49. Trombetta, J. J.; Gennert, D.; Lu, D.; Satija, R.; Shalek, A. K.; Regev, A., Preparation of Single-Cell RNA-Seq Libraries for Next Generation Sequencing. *Curr Protoc Mol Biol* **2014,** *107*, 4 22 1-17.
50. Joensson, H. N.; Andersson Svahn, H., Droplet microfluidics--a tool for single-cell analysis. *Angew Chem Int Ed Engl* **2012,** *51* (49), 12176-92.
51. Ramani, V.; Deng, X.; Qiu, R.; Gunderson, K. L.; Steemers, F. J.; Disteche, C. M.; Noble, W. S.; Duan, Z.; Shendure, J., Massively multiplex single-cell Hi-C. *Nat Methods* **2017,** *14* (3), 263-266.
52. Cao, J.; Packer, J. S.; Ramani, V.; Cusanovich, D. A.; Huynh, C.; Daza, R.; Qiu, X.; Lee, C.; Furlan, S. N.; Steemers, F. J.; Adey, A.; Waterston, R. H.; Trapnell, C.; Shendure, J., Comprehensive single-cell transcriptional profiling of a multicellular organism. *Science* **2017,** *357* (6352), 661-667.
53. Pascal, L. E.; True, L. D.; Campbell, D. S.; Deutsch, E. W.; Risk, M.; Coleman, I. M.; Eichner, L. J.; Nelson, P. S.; Liu, A. Y., Correlation of mRNA and protein levels: cell type-specific gene expression of cluster designation antigens in the prostate. *BMC Genomics* **2008,** *9*, 246.
54. Liu, Y.; Beyer, A.; Aebersold, R., On the Dependency of Cellular Protein Levels on mRNA Abundance. *Cell* **2016,** *165* (3), 535-50.
55. Peterson, V. M.; Zhang, K. X.; Kumar, N.; Wong, J.; Li, L.; Wilson, D. C.; Moore, R.; McClanahan, T. K.; Sadekova, S.; Klappenbach, J. A., Multiplexed quantification of proteins and transcripts in single cells. *Nat Biotechnol* **2017,** *35* (10), 936-939.
56. Baron, M.; Yanai, I., New skin for the old RNA-Seq ceremony: the age of single-cell multi-omics. *Genome Biol* **2017,** *18* (1), 159.
57. Stahlberg, A.; Thomsen, C.; Ruff, D.; Aman, P., Quantitative PCR analysis of DNA, RNAs, and proteins in the same single cell. *Clin Chem* **2012,** *58* (12), 1682-91.
58. Lee, J.; Geiss, G. K.; Demirkan, G.; Vellano, C. P.; Filanoski, B.; Lu, Y.; Ju, Z.; Yu, S.; Guo, H.; Bogatzki, L. Y.; Carter, W. E.; Meredith, R. K.; Krishnamurthy, S.; Ding, Z.; Beechem, J. M.; Mills, G., Implementation of a Multiplex and Quantitative Proteomics Platform for Assessing Protein Lysates Using DNA-Barcoded Antibodies. *Mol Cell Proteomics* **2018**.
59. Genshaft, A. S.; Li, S.; Gallant, C. J.; Darmanis, S.; Prakadan, S. M.; Ziegler, C. G.; Lundberg, M.; Fredriksson, S.; Hong, J.; Regev, A.; Livak, K. J.; Landegren, U.; Shalek, A. K., Multiplexed, targeted profiling of single-cell proteomes and transcriptomes in a single reaction. *Genome Biol* **2016,** *17* (1), 188.
60. Frei, A. P.; Bava, F. A.; Zunder, E. R.; Hsieh, E. W.; Chen, S. Y.; Nolan, G. P.; Gherardini, P. F., Highly multiplexed simultaneous detection of RNAs and proteins in single cells. *Nat Methods* **2016,** *13* (3), 269-75.
61. Bendall, S. C.; Nolan, G. P.; Roederer, M.; Chattopadhyay, P. K., A deep profiler's guide to cytometry. *Trends Immunol* **2012,** *33* (7), 323-32.
62. Stoeckius, M.; Zheng, S.; Houck-Loomis, B.; Hao, S.; Yeung, B.; Smibert, P.; Satija, R., Cell "hashing" with barcoded antibodies enables multiplexing and doublet detection for single cell genomics. *bioRxiv* **2017**.
63. Nag, S.; Dalgaard, M. D.; Kofoed, P. E.; Ursing, J.; Crespo, M.; Andersen, L. O.; Aarestrup, F. M.; Lund, O.; Alifrangis, M., High throughput resistance profiling of Plasmodium falciparum infections based on custom dual indexing and Illumina next generation sequencing-technology. *Sci Rep* **2017,** *7* (1), 2398.

64. Delley, C. L.; Liu, L.; Sarhan, M. F.; Abate, A. R., Combined aptamer and transcriptome sequencing of single cells. *Sci Rep* **2018,** *8* (1), 2919.
65. Ellington, A. D.; Szostak, J. W., In vitro selection of RNA molecules that bind specific ligands. *Nature* **1990,** *346* (6287), 818-22.
66. Delac, M.; Motaln, H.; Ulrich, H.; Lah, T. T., Aptamer for imaging and therapeutic targeting of brain tumor glioblastoma. *Cytometry A* **2015,** *87* (9), 806-16.
67. Dunn, M. R.; Jimenez, R. M.; Chaput, J. C., Analysis of aptamer discovery and technology. *Nature Reviews Chemistry* **2017,** *1*, 0076.
68. Wang, J.; Yu, J.; Yang, Q.; McDermott, J.; Scott, A.; Vukovich, M.; Lagrois, R.; Gong, Q.; Greenleaf, W.; Eisenstein, M.; Ferguson, B. S.; Soh, H. T., Multiparameter Particle Display (MPPD): A Quantitative Screening Method for the Discovery of Highly Specific Aptamers. *Angew Chem Int Ed Engl* **2017,** *56* (3), 744-747.
69. Chen, A.; Yang, S., Replacing antibodies with aptamers in lateral flow immunoassay. *Biosens Bioelectron* **2015,** *71*, 230-242.
70. Lyons, E.; Sheridan, P.; Tremmel, G.; Miyano, S.; Sugano, S., Large-scale DNA Barcode Library Generation for Biomolecule Identification in High-throughput Screens. *Sci Rep* **2017,** *7* (1), 13899.
71. Carvalho, A. M.; Manicardi, A.; Montes, C. V.; Gunnoo, S. B.; Schneider, R. J.; Madder, A., Decoration of trastuzumab with short oligonucleotides: synthesis and detailed characterization. *Org Biomol Chem* **2017,** *15* (42), 8923-8928.
72. Lovendahl, K. N.; Hayward, A. N.; Gordon, W. R., Sequence-Directed Covalent Protein-DNA Linkages in a Single Step Using HUH-Tags. *J Am Chem Soc* **2017,** *139* (20), 7030-7035.
73. Grecco, H. E.; Imtiaz, S.; Zamir, E., Multiplexed imaging of intracellular protein networks. *Cytometry A* **2016,** *89* (8), 761-75.
74. Cheng, Y.; Newell, E. W., Deep Profiling Human T Cell Heterogeneity by Mass Cytometry. *Adv Immunol* **2016,** *131*, 101-34.
75. Bentzen, A. K.; Marquard, A. M.; Lyngaa, R.; Saini, S. K.; Ramskov, S.; Donia, M.; Such, L.; Furness, A. J.; McGranahan, N.; Rosenthal, R.; Straten, P. T.; Szallasi, Z.; Svane, I. M.; Swanton, C.; Quezada, S. A.; Jakobsen, S. N.; Eklund, A. C.; Hadrup, S. R., Large-scale detection of antigen-specific T cells using peptide-MHC-I multimers labeled with DNA barcodes. *Nat Biotechnol* **2016,** *34* (10), 1037-1045.
76. Lu, M.; Chan, B. M.; Schow, P. W.; Chang, W. S.; King, C. T., High-throughput screening of hybridoma supernatants using multiplexed fluorescent cell barcoding on live cells. *J Immunol Methods* **2017,** *451*, 20-27.
77. Zimmermann, G.; Neri, D., DNA-encoded chemical libraries: foundations and applications in lead discovery. *Drug Discov Today* **2016,** *21* (11), 1828-1834.
78. Franzini, R. M.; Neri, D.; Scheuermann, J., DNA-encoded chemical libraries: advancing beyond conventional small-molecule libraries. *Acc Chem Res* **2014,** *47* (4), 1247-55.
79. Yachie, N.; Petsalaki, E.; Mellor, J. C.; Weile, J.; Jacob, Y.; Verby, M.; Ozturk, S. B.; Li, S.; Cote, A. G.; Mosca, R.; Knapp, J. J.; Ko, M.; Yu, A.; Gebbia, M.; Sahni, N.; Yi, S.; Tyagi, T.; Sheykhkarimli, D.; Roth, J. F.; Wong, C.; Musa, L.; Snider, J.; Liu, Y. C.; Yu, H.; Braun, P.; Stagljar, I.; Hao, T.; Calderwood, M. A.; Pelletier, L.; Aloy, P.; Hill, D. E.; Vidal, M.; Roth, F. P., Pooled-matrix protein interaction screens using Barcode Fusion Genetics. *Mol Syst Biol* **2016,** *12* (4), 863.
80. Pollock, S. B.; Hu, A.; Mou, Y.; Martinko, A. J.; Julien, O.; Hornsby, M.; Ploder, L.; Adams, J. J.; Geng, H.; Muschen, M.; Sidhu, S. S.; Moffat, J.; Wells, J. A., Highly multiplexed

and quantitative cell-surface protein profiling using genetically barcoded antibodies. *Proc Natl Acad Sci U S A* **2018,** *115* (11), 2836-2841.
81. Blundell, J. R.; Levy, S. F., Beyond genome sequencing: lineage tracking with barcodes to study the dynamics of evolution, infection, and cancer. *Genomics* **2014,** *104* (6 Pt A), 417-30.
82. Martin, C. J.; Cadena, A. M.; Leung, V. W.; Lin, P. L.; Maiello, P.; Hicks, N.; Chase, M. R.; Flynn, J. L.; Fortune, S. M., Digitally Barcoding Mycobacterium tuberculosis Reveals In Vivo Infection Dynamics in the Macaque Model of Tuberculosis. *MBio* **2017,** *8* (3).
83. Levy, S. F.; Blundell, J. R.; Venkataram, S.; Petrov, D. A.; Fisher, D. S.; Sherlock, G., Quantitative evolutionary dynamics using high-resolution lineage tracking. *Nature* **2015,** *519* (7542), 181-186.
84. Gresham, D., Evolution: Fitness tracking for adapting populations. *Nature* **2015,** *519* (7542), 164-5.
85. Woodworth, M. B.; Girskis, K. M.; Walsh, C. A., Building a lineage from single cells: genetic techniques for cell lineage tracking. *Nat Rev Genet* **2017,** *18* (4), 230-244.
86. Kalhor, R.; Mali, P.; Church, G. M., Rapidly evolving homing CRISPR barcodes. *Nat Methods* **2017,** *14* (2), 195-200.
87. Schmierer, B.; Botla, S. K.; Zhang, J.; Turunen, M.; Kivioja, T.; Taipale, J., CRISPR/Cas9 screening using unique molecular identifiers. *Molecular Systems Biology* **2017,** *13* (10), 945.
88. McKenna, A.; Findlay, G. M.; Gagnon, J. A.; Horwitz, M. S.; Schier, A. F.; Shendure, J., Whole-organism lineage tracing by combinatorial and cumulative genome editing. *Science* **2016,** *353* (6298), aaf7907.
89. Raj, B.; Wagner, D. E.; McKenna, A.; Pandey, S.; Klein, A. M.; Shendure, J.; Gagnon, J. A.; Schier, A. F., Simultaneous single-cell profiling of lineages and cell types in the vertebrate brain. *Nat Biotechnol* **2018**.
90. Adamson, B.; Norman, T. M.; Jost, M.; Cho, M. Y.; Nunez, J. K.; Chen, Y.; Villalta, J. E.; Gilbert, L. A.; Horlbeck, M. A.; Hein, M. Y.; Pak, R. A.; Gray, A. N.; Gross, C. A.; Dixit, A.; Parnas, O.; Regev, A.; Weissman, J. S., A Multiplexed Single-Cell CRISPR Screening Platform Enables Systematic Dissection of the Unfolded Protein Response. *Cell* **2016,** *167* (7), 1867-1882 e21.
91. Dixit, A.; Parnas, O.; Li, B.; Chen, J.; Fulco, C. P.; Jerby-Arnon, L.; Marjanovic, N. D.; Dionne, D.; Burks, T.; Raychowdhury, R.; Adamson, B.; Norman, T. M.; Lander, E. S.; Weissman, J. S.; Friedman, N.; Regev, A., Perturb-Seq: Dissecting Molecular Circuits with Scalable Single-Cell RNA Profiling of Pooled Genetic Screens. *Cell* **2016,** *167* (7), 1853-1866 e17.
92. Datlinger, P.; Rendeiro, A. F.; Schmidl, C.; Krausgruber, T.; Traxler, P.; Klughammer, J.; Schuster, L. C.; Kuchler, A.; Alpar, D.; Bock, C., Pooled CRISPR screening with single-cell transcriptome readout. *Nat Methods* **2017,** *14* (3), 297-301.
93. Fennessey, C. M.; Pinkevych, M.; Immonen, T. T.; Reynaldi, A.; Venturi, V.; Nadella, P.; Reid, C.; Newman, L.; Lipkey, L.; Oswald, K.; Bosche, W. J.; Trivett, M. T.; Ohlen, C.; Ott, D. E.; Estes, J. D.; Del Prete, G. Q.; Lifson, J. D.; Davenport, M. P.; Keele, B. F., Genetically-barcoded SIV facilitates enumeration of rebound variants and estimation of reactivation rates in nonhuman primates following interruption of suppressive antiretroviral therapy. *PLoS Pathog* **2017,** *13* (5), e1006359.

94. Livet, J.; Weissman, T. A.; Kang, H.; Draft, R. W.; Lu, J.; Bennis, R. A.; Sanes, J. R.; Lichtman, J. W., Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system. *Nature* **2007,** *450* (7166), 56-62.
95. Zador, A. M.; Dubnau, J.; Oyibo, H. K.; Zhan, H.; Cao, G.; Peikon, I. D., Sequencing the connectome. *PLoS Biol* **2012,** *10* (10), e1001411.
96. Xiong, F.; Obholzer, N. D.; Noche, R. R.; Megason, S. G., Multibow: digital spectral barcodes for cell tracing. *PLoS One* **2015,** *10* (5), e0127822.
97. Nilsson, M.; Malmgren, H.; Samiotaki, M.; Kwiatkowski, M.; Chowdhary, B. P.; Landegren, U., Padlock probes: circularizing oligonucleotides for localized DNA detection. *Science* **1994,** *265* (5181), 2085-8.
98. Ke, R.; Mignardi, M.; Pacureanu, A.; Svedlund, J.; Botling, J.; Wahlby, C.; Nilsson, M., In situ sequencing for RNA analysis in preserved tissue and cells. *Nat Methods* **2013,** *10* (9), 857-60.
99. Larsson, C.; Koch, J.; Nygren, A.; Janssen, G.; Raap, A. K.; Landegren, U.; Nilsson, M., In situ genotyping individual DNA molecules by target-primed rolling-circle amplification of padlock probes. *Nat Methods* **2004,** *1* (3), 227-32.
100. Lubeck, E.; Coskun, A. F.; Zhiyentayev, T.; Ahmad, M.; Cai, L., Single cell in situ RNA profiling by sequential hybridization. *Nature methods* **2014,** *11* (4), 360-361.
101. Cai, L., Turning single cells into microarrays by super-resolution barcoding. *Briefings in Functional Genomics* **2013,** *12* (2), 75-80.
102. Tang, S.; Zhang, Y.; Dhakal, P.; Ravelo, L.; Anderson, C. L.; Collins, K. M.; Raymo, F. M., Photochemical Barcodes. *J Am Chem Soc* **2018,** *140* (13), 4485-4488.
103. Han, M.; Gao, X.; Su, J. Z.; Nie, S., Quantum-dot-tagged microbeads for multiplexed optical coding of biomolecules. *Nat Biotechnol* **2001,** *19* (7), 631-5.
104. Levsky, J. M.; Shenoy, S. M.; Pezo, R. C.; Singer, R. H., Single-Cell Gene Expression Profiling. *Science* **2002,** *297* (5582), 836.
105. Braeckmans, K.; De Smedt, S. C.; Roelant, C.; Leblans, M.; Pauwels, R.; Demeester, J., Encoding microcarriers by spatial selective photobleaching. *Nat Mater* **2003,** *2* (3), 169-73.
106. Geiss, G. K.; Bumgarner, R. E.; Birditt, B.; Dahl, T.; Dowidar, N.; Dunaway, D. L.; Fell, H. P.; Ferree, S.; George, R. D.; Grogan, T.; James, J. J.; Maysuria, M.; Mitton, J. D.; Oliveri, P.; Osborn, J. L.; Peng, T.; Ratcliffe, A. L.; Webster, P. J.; Davidson, E. H.; Hood, L.; Dimitrov, K., Direct multiplexed measurement of gene expression with color-coded probe pairs. *Nat Biotechnol* **2008,** *26* (3), 317-25.
107. Lin, C.; Jungmann, R.; Leifer, A. M.; Li, C.; Levner, D.; Church, G. M.; Shih, W. M.; Yin, P., Submicrometre geometrically encoded fluorescent barcodes self-assembled from DNA. *Nat Chem* **2012,** *4* (10), 832-9.
108. Nicewarner-Pena, S. R.; Freeman, R. G.; Reiss, B. D.; He, L.; Pena, D. J.; Walton, I. D.; Cromer, R.; Keating, C. D.; Natan, M. J., Submicrometer metallic barcodes. *Science* **2001,** *294* (5540), 137-41.
109. Hu, F.; Zeng, C.; Long, R.; Miao, Y.; Wei, L.; Xu, Q.; Min, W., Supermultiplexed optical imaging and barcoding with engineered polyynes. *Nat Methods* **2018,** *15* (3), 194-200.
110. Angelo, M.; Bendall, S. C.; Finck, R.; Hale, M. B.; Hitzman, C.; Borowsky, A. D.; Levenson, R. M.; Lowe, J. B.; Liu, S. D.; Zhao, S.; Natkunam, Y.; Nolan, G. P., Multiplexed ion beam imaging of human breast tumors. *Nat Med* **2014,** *20* (4), 436-42.
111. Keren, L.; Bosse, M.; Marquez, D.; Angoshtari, R.; Jain, S.; Varma, S.; Yang, S. R.; Kurian, A.; Van Valen, D.; West, R.; Bendall, S. C.; Angelo, M., A Structured Tumor-Immune

Microenvironment in Triple Negative Breast Cancer Revealed by Multiplexed Ion Beam Imaging. *Cell* **2018,** *174* (6), 1373-1387 e19.
112. Levsky, J. M.; Singer, R. H., Fluorescence in situ hybridization: past, present and future. *J Cell Sci* **2003,** *116* (Pt 14), 2833-8.
113. Satija, R.; Farrell, J. A.; Gennert, D.; Schier, A. F.; Regev, A., Spatial reconstruction of single-cell gene expression data. *Nat Biotechnol* **2015,** *33* (5), 495-502.
114. Achim, K.; Pettit, J. B.; Saraiva, L. R.; Gavriouchkina, D.; Larsson, T.; Arendt, D.; Marioni, J. C., High-throughput spatial mapping of single-cell RNA-seq data to tissue of origin. *Nat Biotechnol* **2015,** *33* (5), 503-9.
115. Faridani, O. R.; Sandberg, R., Putting cells in their place. *Nat Biotechnol* **2015,** *33* (5), 490-1.
116. Lee, J. H.; Daugharthy, E. R.; Scheiman, J.; Kalhor, R.; Ferrante, T. C.; Terry, R.; Turczyk, B. M.; Yang, J. L.; Lee, H. S.; Aach, J.; Zhang, K.; Church, G. M., Fluorescent in situ sequencing (FISSEQ) of RNA for gene expression profiling in intact cells and tissues. *Nat Protoc* **2015,** *10* (3), 442-58.
117. Medaglia, C.; Giladi, A.; Stoler-Barak, L.; De Giovanni, M.; Salame, T. M.; Biram, A.; David, E.; Li, H.; Iannacone, M.; Shulman, Z.; Amit, I., Spatial reconstruction of immune niches by combining photoactivatable reporters and scRNA-seq. *Science* **2017,** *358* (6370), 1622-1626.
118. Victora, G. D.; Schwickert, T. A.; Fooksman, D. R.; Kamphorst, A. O.; Meyer-Hermann, M.; Dustin, M. L.; Nussenzweig, M. C., Germinal center dynamics revealed by multiphoton microscopy with a photoactivatable fluorescent reporter. *Cell* **2010,** *143* (4), 592-605.
119. Chtanova, T.; Hampton, H. R.; Waterhouse, L. A.; Wood, K.; Tomura, M.; Miwa, Y.; Mackay, C. R.; Brink, R.; Phan, T. G., Real-time interactive two-photon photoconversion of recirculating lymphocytes for discontinuous cell tracking in live adult mice. *J Biophotonics* **2014,** *7* (6), 425-33.
120. Moran, I.; Phan, T. G., Fate Mapping and Transcript Profiling of Germinal Center Cells by Two-Photon Photoconversion. *Methods Mol Biol* **2017,** *1623*, 59-72.
121. Stahl, P. L.; Salmen, F.; Vickovic, S.; Lundmark, A.; Navarro, J. F.; Magnusson, J.; Giacomello, S.; Asp, M.; Westholm, J. O.; Huss, M.; Mollbrink, A.; Linnarsson, S.; Codeluppi, S.; Borg, A.; Ponten, F.; Costea, P. I.; Sahlen, P.; Mulder, J.; Bergmann, O.; Lundeberg, J.; Frisen, J., Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science* **2016,** *353* (6294), 78-82.
122. Belisle, J. M.; Mazzaferri, J.; Costantino, S., Laser-assisted adsorption by photobleaching. *Methods Cell Biol* **2014,** *119*, 125-40.
123. Belisle, J. M.; Kunik, D.; Costantino, S., Rapid multicomponent optical protein patterning. *Lab Chip* **2009,** *9* (24), 3580-5.
124. Chilkoti, A.; Stayton, P. S., Molecular Origins of the Slow Streptavidin-Biotin Dissociation Kinetics. *Journal of the American Chemical Society* **1995,** *117* (43), 10622-10628.
125. Sanderson, M. J.; Smith, I.; Parker, I.; Bootman, M. D., Fluorescence microscopy. *Cold Spring Harb Protoc* **2014,** *2014* (10), pdb top071795.
126. Song, L.; Hennink, E. J.; Young, I. T.; Tanke, H. J., Photobleaching kinetics of fluorescein in quantitative fluorescence microscopy. *Biophys J* **1995,** *68* (6), 2588-600.
127. Phaniendra, A.; Jestadi, D. B.; Periyasamy, L., Free radicals: properties, sources, targets, and their implication in various diseases. *Indian J Clin Biochem* **2015,** *30* (1), 11-26.
128. Cheeseman, K. H.; Slater, T. F., An introduction to free radical biochemistry. *Br Med Bull* **1993,** *49* (3), 481-93.

129. Petersen, R. C.; Reddy, M. S.; Liu, P. R., Advancements in Free-Radical Pathologies and an Important Treatment Solution with a Free-Radical Inhibitor. *SF J Biotechnol Biomed Eng* **2018,** *1* (1).
130. Mason, R. P., Imaging free radicals in organelles, cells, tissue, and in vivo with immuno-spin trapping. *Redox Biol* **2016,** *8*, 422-9.
131. Ma, Q.; Shen, J. H.; Shen, S. R.; Das, U. N., Bioactive lipids in pathological retinopathy. *Crit Rev Food Sci Nutr* **2014,** *54* (1), 1-16.
132. Kukielka, E.; Puntarulo, S.; Cederbaum, A. I., Interaction of ferric complexes with rat liver nuclei to catalyze NADH-and NADPH-Dependent production of oxygen radicals. *Arch Biochem Biophys* **1989,** *273* (2), 319-30.
133. Wang, M.; Orwar, O.; Olofsson, J.; Weber, S. G., Single-cell electroporation. *Anal Bioanal Chem* **2010,** *397* (8), 3235-48.
134. Teissie, J.; Golzio, M.; Rols, M. P., Mechanisms of cell membrane electropermeabilization: a minireview of our present (lack of ?) knowledge. *Biochimica et biophysica acta* **2005,** *1724* (3), 270-80.
135. Zhang, Y.; Yu, L. C., Single-cell microinjection technology in cell biology. *Bioessays* **2008,** *30* (6), 606-10.
136. Tsulaia, T. V.; Prokopishyn, N. L.; Yao, A.; Carsrud, N. D.; Carou, M. C.; Brown, D. B.; Davis, B. R.; Yannariello-Brown, J., Glass needle-mediated microinjection of macromolecules and transgenes into primary human mesenchymal stem cells. *J Biomed Sci* **2003,** *10* (3), 328-36.
137. Espina, V.; Wulfkuhle, J. D.; Calvert, V. S.; VanMeter, A.; Zhou, W.; Coukos, G.; Geho, D. H.; Petricoin, E. F., 3rd; Liotta, L. A., Laser-capture microdissection. *Nature protocols* **2006,** *1* (2), 586-603.
138. Zhou, X. X.; Lin, M. Z., Photoswitchable fluorescent proteins: ten years of colorful chemistry and exciting applications. *Curr Opin Chem Biol* **2013,** *17* (4), 682-90.
139. Dempsey, W. P.; Georgieva, L.; Helbling, P. M.; Sonay, A. Y.; Truong, T. V.; Haffner, M.; Pantazis, P., In vivo single-cell labeling by confined primed conversion. *Nature methods* **2015,** *12* (7), 645-8.
140. Kanter, I.; Kalisky, T., Single cell transcriptomics: methods and applications. *Frontiers in oncology* **2015,** *5*, 53.
141. Wang, Y.; Navin, N. E., Advances and applications of single-cell sequencing technologies. *Molecular cell* **2015,** *58* (4), 598-609.
142. Belisle, J. M.; Correia, J. P.; Wiseman, P. W.; Kennedy, T. E.; Costantino, S., Patterning protein concentration using laser-assisted adsorption by photobleaching, LAPAP. *Lab Chip* **2008,** *8* (12), 2164-7.
143. Chien, M. P.; Werley, C. A.; Farhi, S. L.; Cohen, A. E., Photostick: a method for selective isolation of target cells from culture. *Chem Sci* **2015,** *6* (3), 1701-1705.
144. Kolodziejczyk, A. A.; Kim, J. K.; Svensson, V.; Marioni, J. C.; Teichmann, S. A., The technology and biology of single-cell RNA sequencing. *Molecular cell* **2015,** *58* (4), 610-20.
145. Otsu, N., Threshold Selection Method from Gray-Level Histograms. *Ieee T Syst Man Cyb* **1979,** *9* (1), 62-66.
146. Doering, L. C., *Protocols for neural cell culture*. 4th ed.; Humana Press: New York, 2010; p xi, 400 p.
147. Bolger, A. M.; Lohse, M.; Usadel, B., Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **2014,** *30* (15), 2114-20.

148. Quinlan, A. R.; Hall, I. M., BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **2010,** *26* (6), 841-2.
149. Wang, L.; Wang, S.; Li, W., RSeQC: quality control of RNA-seq experiments. *Bioinformatics* **2012,** *28* (16), 2184-5.
150. Thorvaldsdottir, H.; Robinson, J. T.; Mesirov, J. P., Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief Bioinform* **2013,** *14* (2), 178-92.
151. Li, H.; Handsaker, B.; Wysoker, A.; Fennell, T.; Ruan, J.; Homer, N.; Marth, G.; Abecasis, G.; Durbin, R.; Genome Project Data Processing, S., The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **2009,** *25* (16), 2078-9.
152. Genomes Project, C.; Abecasis, G. R.; Altshuler, D.; Auton, A.; Brooks, L. D.; Durbin, R. M.; Gibbs, R. A.; Hurles, M. E.; McVean, G. A., A map of human genome variation from population-scale sequencing. *Nature* **2010,** *467* (7319), 1061-73.
153. Patterson, D. M.; Prescher, J. A., Orthogonal bioorthogonal chemistries. *Curr Opin Chem Biol* **2015,** *28*, 141-9.
154. Braun, A. C.; Gutmann, M.; Luhmann, T.; Meinel, L., Bioorthogonal strategies for site-directed decoration of biomaterials with therapeutic proteins. *J Control Release* **2018,** *273*, 68-85.
155. Gautier, A.; Juillerat, A.; Heinis, C.; Correa, I. R., Jr.; Kindermann, M.; Beaufils, F.; Johnsson, K., An engineered protein tag for multiprotein labeling in living cells. *Chem Biol* **2008,** *15* (2), 128-36.
156. Pappas, D.; Wang, K., Cellular separations: a review of new challenges in analytical chemistry. *Anal Chim Acta* **2007,** *601* (1), 26-35.
157. Kuka, M.; Ashwell, J. D., A method for high purity sorting of rare cell subsets applied to TDC. *J Immunol Methods* **2013,** *400-401*, 111-6.
158. Kang, J. H.; Krause, S.; Tobin, H.; Mammoto, A.; Kanapathipillai, M.; Ingber, D. E., A combined micromagnetic-microfluidic device for rapid capture and culture of rare circulating tumor cells. *Lab Chip* **2012,** *12* (12), 2175-81.
159. Miltenyi, S.; Muller, W.; Weichel, W.; Radbruch, A., High gradient magnetic cell separation with MACS. *Cytometry* **1990,** *11* (2), 231-8.
160. Shields, C. W. t.; Reyes, C. D.; Lopez, G. P., Microfluidic cell sorting: a review of the advances in the separation of cells from debulking to rare cell isolation. *Lab Chip* **2015,** *15* (5), 1230-49.
161. Takahashi, K.; Okada, T. S., An analysis of the effect of "conditioned medium" upon the cell culture at low density. *Dev Growth Differ* **1970,** *12* (2), 65-77.
162. Huang, B.; Bird, S.; Kemble, R.; Simmonds, D.; Keller, W.; Miki, B., Effects of culture density, conditioned medium and feeder cultures on microspore embryogenesis in Brassica napus L. cv. Topas. *Plant Cell Rep* **1990,** *8* (10), 594-7.
163. Housden, B. E.; Valvezan, A. J.; Kelley, C.; Sopko, R.; Hu, Y.; Roesel, C.; Lin, S.; Buckner, M.; Tao, R.; Yilmazel, B.; Mohr, S. E.; Manning, B. D.; Perrimon, N., Identification of potential drug targets for tuberous sclerosis complex by synthetic screens combining CRISPR-based knockouts with RNAi. *Sci Signal* **2015,** *8* (393), rs9.
164. Yamamoto, A.; Mishima, S.; Maruyama, N.; Sumita, M., Quantitative evaluation of cell attachment to glass, polystyrene, and fibronectin- or collagen-coated polystyrene by measurement of cell adhesive shear force and cell detachment energy. *J Biomed Mater Res* **2000,** *50* (2), 114-24.

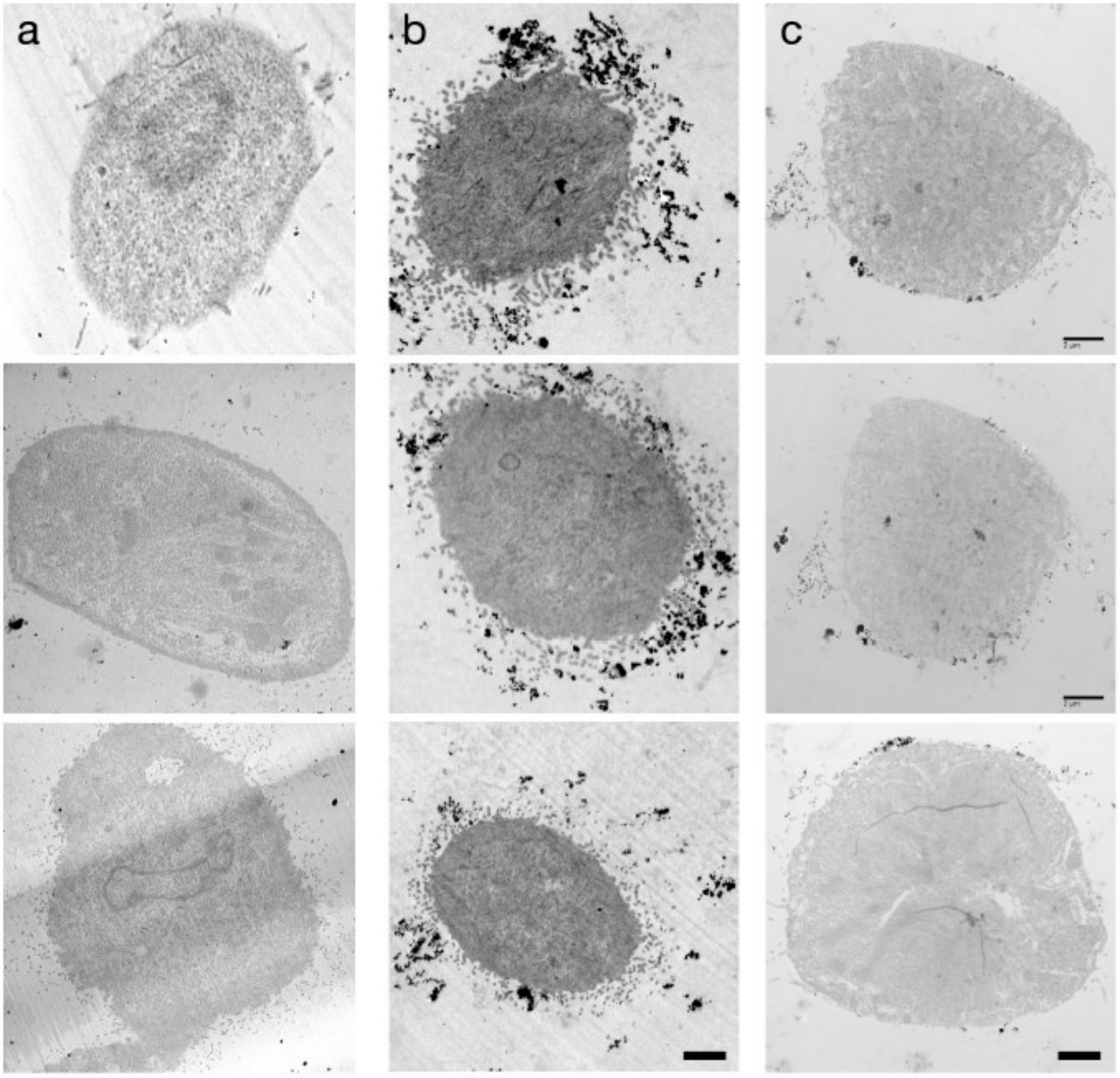
165. Fradet-Turcotte, A.; Canny, M. D.; Escribano-Diaz, C.; Orthwein, A.; Leung, C. C.; Huang, H.; Landry, M. C.; Kitevski-LeBlanc, J.; Noordermeer, S. M.; Sicheri, F.; Durocher, D., 53BP1 is a reader of the DNA-damage-induced H2A Lys 15 ubiquitin mark. *Nature* **2013,** *499* (7456), 50-4.
166. Asaithamby, A.; Chen, D. J., Cellular responses to DNA double-strand breaks after low-dose gamma-irradiation. *Nucleic Acids Res* **2009,** *37* (12), 3912-23.
167. Mirzayans, R.; Andrais, B.; Murray, D., Roles of Polyploid/Multinucleated Giant Cancer Cells in Metastasis and Disease Relapse Following Anticancer Treatment. *Cancers (Basel)* **2018,** *10* (4).
168. Mirzayans, R.; Andrais, B.; Scott, A.; Wang, Y. W.; Kumar, P.; Murray, D., Multinucleated Giant Cancer Cells Produced in Response to Ionizing Radiation Retain Viability and Replicate Their Genome. *Int J Mol Sci* **2017,** *18* (2).
169. Coward, J.; Harding, A., Size Does Matter: Why Polyploid Tumor Cells are Critical Drug Targets in the War on Cancer. *Front Oncol* **2014,** *4*, 123.
170. Mittal, K.; Donthamsetty, S.; Kaur, R.; Yang, C.; Gupta, M. V.; Reid, M. D.; Choi, D. H.; Rida, P. C. G.; Aneja, R., Multinucleated polyploidy drives resistance to Docetaxel chemotherapy in prostate cancer. *Br J Cancer* **2017,** *116* (9), 1186-1194.
171. Weihua, Z.; Lin, Q.; Ramoth, A. J.; Fan, D.; Fidler, I. J., Formation of solid tumors by a single multinucleated cancer cell. *Cancer* **2011,** *117* (17), 4092-9.
172. Green, H.; Meuth, M., An established pre-adipose cell line and its differentiation in culture. *Cell* **1974,** *3* (2), 127-33.
173. Armani, A.; Mammi, C.; Marzolla, V.; Calanchini, M.; Antelmi, A.; Rosano, G. M.; Fabbri, A.; Caprio, M., Cellular models for understanding adipogenesis, adipose dysfunction, and obesity. *J Cell Biochem* **2010,** *110* (3), 564-72.
174. Majka, S. M.; Miller, H. L.; Helm, K. M.; Acosta, A. S.; Childs, C. R.; Kong, R.; Klemm, D. J., Analysis and isolation of adipocytes by flow cytometry. *Methods Enzymol* **2014,** *537*, 281-96.
175. Antfolk, M.; Kim, S. H.; Koizumi, S.; Fujii, T.; Laurell, T., Label-free single-cell separation and imaging of cancer cells using an integrated microfluidic system. *Sci Rep* **2017,** *7*, 46507.
176. Moon, H. S.; Kwon, K.; Kim, S. I.; Han, H.; Sohn, J.; Lee, S.; Jung, H. I., Continuous separation of breast cancer cells from blood samples using multi-orifice flow fractionation (MOFF) and dielectrophoresis (DEP). *Lab Chip* **2011,** *11* (6), 1118-25.
177. Tham, Y. C.; Li, X.; Wong, T. Y.; Quigley, H. A.; Aung, T.; Cheng, C. Y., Global Prevalence of Glaucoma and Projections of Glaucoma Burden through 2040 A Systematic Review and Meta-Analysis. *Ophthalmology* **2014,** *121* (11), 2081-2090.
178. Pamme, N.; Wilhelm, C., Continuous sorting of magnetic cells via on-chip free-flow magnetophoresis. *Lab Chip* **2006,** *6* (8), 974-80.
179. Radbruch, A.; Mechtold, B.; Thiel, A.; Miltenyi, S.; Pflüger, E., Chapter 23 High-Gradient Magnetic Cell Sorting. In *Methods in Cell Biology*, Darzynkiewicz, Z.; Robinson, J. P.; Crissman, H. A., Eds. Academic Press: 1994; Vol. 42, pp 387-403.
180. Khojah, R.; Stoutamore, R.; Di Carlo, D., Size-tunable microvortex capture of rare cells. *Lab Chip* **2017**.
181. Zhao, W.; Cheng, R.; Lim, S. H.; Miller, J. R.; Zhang, W.; Tang, W.; Xie, J.; Mao, L., Biocompatible and label-free separation of cancer cells from cell culture lines from white blood cells in ferrofluids. *Lab Chip* **2017**.

182. Monti, M.; Imberti, B.; Bianchi, N.; Pezzotta, A.; Morigi, M.; Del Fante, C.; Redi, C. A.; Perotti, C., A novel method for the isolation of pluripotent stem cells from human umbilical cord blood. *Stem Cells Dev* **2017**.
183. Hosokawa, M.; Hayata, T.; Fukuda, Y.; Arakaki, A.; Yoshino, T.; Tanaka, T.; Matsunaga, T., Size-selective microcavity array for rapid and efficient detection of circulating tumor cells. *Anal Chem* **2010,** *82* (15), 6629-35.
184. Davis, J. A.; Inglis, D. W.; Morton, K. J.; Lawrence, D. A.; Huang, L. R.; Chou, S. Y.; Sturm, J. C.; Austin, R. H., Deterministic hydrodynamics: taking blood apart. *Proc Natl Acad Sci U S A* **2006,** *103* (40), 14779-84.
185. Landry, J. P.; Ke, Y.; Yu, G. L.; Zhu, X. D., Measuring affinity constants of 1450 monoclonal antibodies to peptide targets with a microarray-based label-free assay platform. *J Immunol Methods* **2015,** *417*, 86-96.
186. Stubenrauch, K.; Wessels, U.; Essig, U.; Kowalewsky, F.; Vogel, R.; Heinrich, J., Characterization of murine anti-human Fab antibodies for use in an immunoassay for generic quantification of human Fab fragments in non-human serum samples including cynomolgus monkey samples. *J Pharm Biomed Anal* **2013,** *72*, 208-15.
187. Landais, E.; Romagnoli, P. A.; Corper, A. L.; Shires, J.; Altman, J. D.; Wilson, I. A.; Garcia, K. C.; Teyton, L., New design of MHC class II tetramers to accommodate fundamental principles of antigen presentation. *J Immunol* **2009,** *183* (12), 7949-57.
188. Wooldridge, L.; Lissina, A.; Cole, D. K.; van den Berg, H. A.; Price, D. A.; Sewell, A. K., Tricks with tetramers: how to get the most from multimeric peptide-MHC. *Immunology* **2009,** *126* (2), 147-164.
189. Tsai, S. L.; Lee, T. H.; Chien, R. N.; Liao, S. K.; Lin, C. L.; Kuo, G. C.; Liaw, Y. F., A method to increase tetramer staining efficiency of CD8+ T cells with MHC-peptide complexes: therapeutic applications in monitoring cytotoxic T lymphocyte activity during hepatitis B and C treatment. *J Immunol Methods* **2004,** *285* (1), 71-87.
190. van der Toom, E. E.; Verdone, J. E.; Jun, C.; Petrisor, D.; Lim, S.; de la Rosette, J. J.; de Reijke, T. M.; Gorin, M. A.; Pienta, K. J.; Stoianovici, D., A surface tension magnetophoretic device for rare cell isolation and characterization. *Med Oncol* **2017,** *34* (2), 22.
191. Navin, N.; Kendall, J.; Troge, J.; Andrews, P.; Rodgers, L.; McIndoo, J.; Cook, K.; Stepansky, A.; Levy, D.; Esposito, D.; Muthuswamy, L.; Krasnitz, A.; McCombie, W. R.; Hicks, J.; Wigler, M., Tumour evolution inferred by single-cell sequencing. *Nature* **2011,** *472* (7341), 90-4.
192. Valastyan, S.; Weinberg, R. A., Tumor metastasis: molecular insights and evolving paradigms. *Cell* **2011,** *147* (2), 275-92.
193. Shapiro, E.; Biezuner, T.; Linnarsson, S., Single-cell sequencing-based technologies will revolutionize whole-organism science. *Nat Rev Genet* **2013,** *14* (9), 618-30.
194. Tirosh, I.; Izar, B.; Prakadan, S. M.; Wadsworth, M. H., 2nd; Treacy, D.; Trombetta, J. J.; Rotem, A.; Rodman, C.; Lian, C.; Murphy, G.; Fallahi-Sichani, M.; Dutton-Regester, K.; Lin, J. R.; Cohen, O.; Shah, P.; Lu, D.; Genshaft, A. S.; Hughes, T. K.; Ziegler, C. G.; Kazer, S. W.; Gaillard, A.; Kolb, K. E.; Villani, A. C.; Johannessen, C. M.; Andreev, A. Y.; Van Allen, E. M.; Bertagnolli, M.; Sorger, P. K.; Sullivan, R. J.; Flaherty, K. T.; Frederick, D. T.; Jane-Valbuena, J.; Yoon, C. H.; Rozenblatt-Rosen, O.; Shalek, A. K.; Regev, A.; Garraway, L. A., Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science* **2016,** *352* (6282), 189-96.

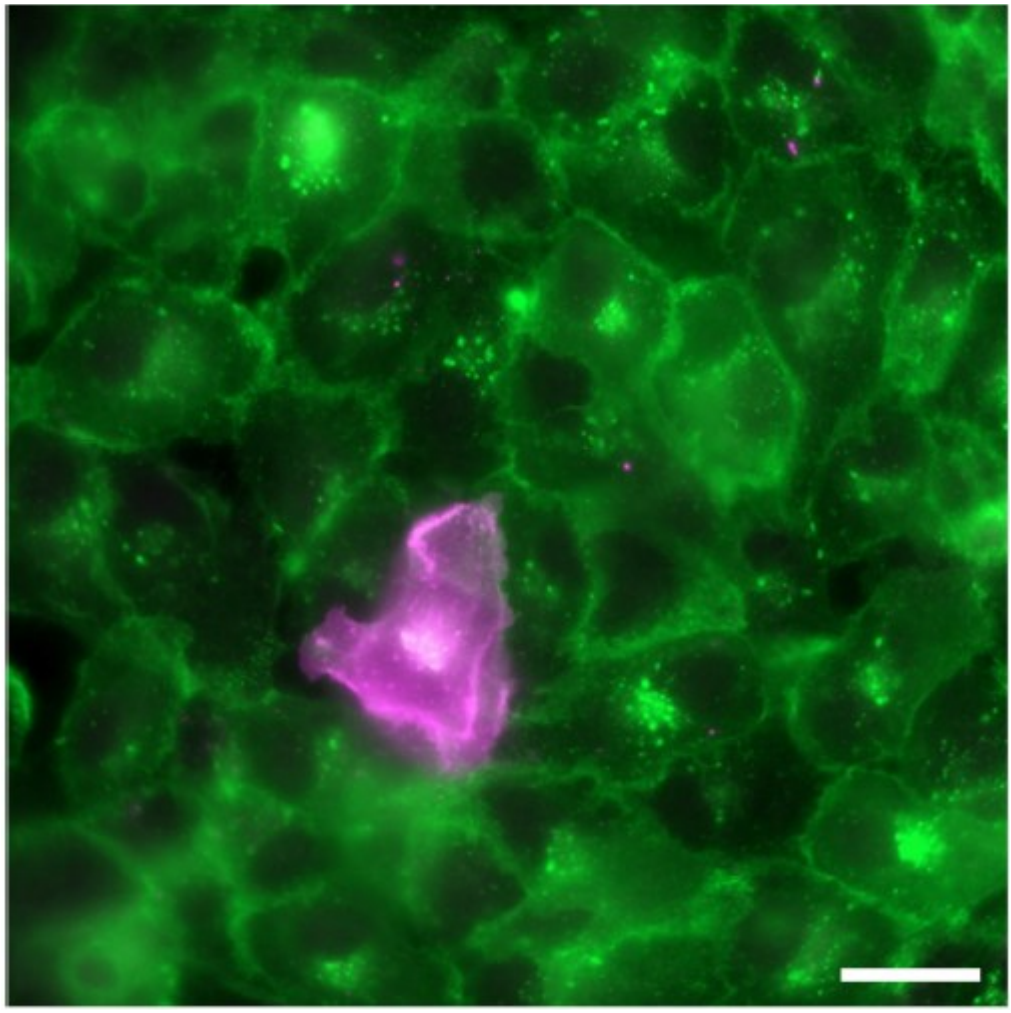
195. Heitzer, E.; Auer, M.; Gasch, C.; Pichler, M.; Ulz, P.; Hoffmann, E. M.; Lax, S.; Waldispuehl-Geigl, J.; Mauermann, O.; Lackner, C.; Hofler, G.; Eisner, F.; Sill, H.; Samonigg, H.; Pantel, K.; Riethdorf, S.; Bauernhofer, T.; Geigl, J. B.; Speicher, M. R., Complex tumor genomes inferred from single circulating tumor cells by array-CGH and next-generation sequencing. *Cancer Res* **2013,** *73* (10), 2965-75.
196. Gierahn, T. M.; Wadsworth, M. H., 2nd; Hughes, T. K.; Bryson, B. D.; Butler, A.; Satija, R.; Fortune, S.; Love, J. C.; Shalek, A. K., Seq-Well: portable, low-cost RNA sequencing of single cells at high throughput. *Nat Methods* **2017,** *14* (4), 395-398.
197. Brennecke, P.; Anders, S.; Kim, J. K.; Kolodziejczyk, A. A.; Zhang, X.; Proserpio, V.; Baying, B.; Benes, V.; Teichmann, S. A.; Marioni, J. C.; Heisler, M. G., Accounting for technical noise in single-cell RNA-seq experiments. *Nat Methods* **2013,** *10* (11), 1093-5.
198. Wu, A. R.; Neff, N. F.; Kalisky, T.; Dalerba, P.; Treutlein, B.; Rothenberg, M. E.; Mburu, F. M.; Mantalas, G. L.; Sim, S.; Clarke, M. F.; Quake, S. R., Quantitative assessment of single-cell RNA-sequencing methods. *Nat Methods* **2014,** *11* (1), 41-6.
199. Tan, S. J.; Yeo, T.; Sukhatme, S. A.; Kong, S. L.; Lim, W. T.; Lim, C. T., Personalized Treatment Through Detection and Monitoring of Genetic Aberrations in Single Circulating Tumor Cells. *Adv Exp Med Biol* **2017,** *994*, 255-273.
200. Hood, L.; Heath, J. R.; Phelps, M. E.; Lin, B., Systems Biology and New Technologies Enable Predictive and Preventative Medicine. *Science* **2004,** *306* (5696), 640-643.
201. Pugia, M.; Magbanua, M. J. M.; Park, J. W., Enrichment and Detection of Circulating Tumor Cells and Other Rare Cell Populations by Microfluidic Filtration. *Adv Exp Med Biol* **2017,** *994*, 119-131.
202. Weissleder, R., CHAPTER 43 - Molecular Imaging as a Paradigm for Genomic and Personalized Medicine. In *Genomic and Personalized Medicine*, Willard, P. D. H. F.; Ginsburg, G. S., Eds. Academic Press: New York, 2009; pp 494-499.
203. Hesketh, G. G.; Youn, J. Y.; Samavarchi-Tehrani, P.; Raught, B.; Gingras, A. C., Parallel Exploration of Interaction Space by BioID and Affinity Purification Coupled to Mass Spectrometry. *Methods Mol Biol* **2017,** *1550*, 115-136.
204. Al-Hakim, A. K.; Bashkurov, M.; Gingras, A. C.; Durocher, D.; Pelletier, L., Interaction proteomics identify NEURL4 and the HECT E3 ligase HERC2 as novel modulators of centrosome architecture. *Mol Cell Proteomics* **2012,** *11* (6), M111 014233.
205. Brown, J. P.; Wei, W.; Sedivy, J. M., Bypass of senescence after disruption of p21CIP1/WAF1 gene in normal diploid human fibroblasts. *Science* **1997,** *277* (5327), 831-4.
206. Talbot, S.; Abdulnour, R. E.; Burkett, P. R.; Lee, S.; Cronin, S. J.; Pascal, M. A.; Laedermann, C.; Foster, S. L.; Tran, J. V.; Lai, N.; Chiu, I. M.; Ghasemlou, N.; DiBiase, M.; Roberson, D.; Von Hehn, C.; Agac, B.; Haworth, O.; Seki, H.; Penninger, J. M.; Kuchroo, V. K.; Bean, B. P.; Levy, B. D.; Woolf, C. J., Silencing Nociceptor Neurons Reduces Allergic Airway Inflammation. *Neuron* **2015,** *87* (2), 341-54.
207. Belanger, F.; Fortier, E.; Dube, M.; Lemay, J. F.; Buisson, R.; Masson, J. Y.; Elsherbiny, A.; Costantino, S.; Carmona, E.; Mes-Masson, A. M.; Wurtele, H.; Drobetsky, E., Replication Protein A Availability during DNA Replication Stress Is a Major Determinant of Cisplatin Resistance in Ovarian Cancer Cells. *Cancer Res* **2018,** *78* (19), 5561-5573.
208. Belanger, F.; Angers, J. P.; Fortier, E.; Hammond-Martel, I.; Costantino, S.; Drobetsky, E.; Wurtele, H., Mutations in Replicative Stress Response Pathways Are Associated with S Phase-specific Defects in Nucleotide Excision Repair. *J Biol Chem* **2016,** *291* (2), 522-37.

209. Otsu, N., A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems, Man, and Cybernetics* **1979,** *9* (1), 62-66.
210. Walen, K. H., The origin of transformed cells. studies of spontaneous and induced cell transformation in cell cultures from marsupials, a snail, and human amniocytes. *Cancer Genet Cytogenet* **2002,** *133* (1), 45-54.
211. Feinstein, R. E.; Nikkila, T., Occurrence of multinucleated giant cells in the appendix of clinically healthy rabbits. *J Comp Pathol* **1988,** *99* (4), 439-47.
212. Anderson, C. A.; Roberts, S.; Zhang, H.; Kelly, C. M.; Kendall, A.; Lee, C.; Gerstenberger, J.; Koenig, A. B.; Kabeche, R.; Gladfelter, A. S., Ploidy variation in multinucleate cells changes under stress. *Mol Biol Cell* **2015,** *26* (6), 1129-40.
213. Zaritsky, A.; Manor, N.; Wolf, L.; Ben-Jacob, E.; Tsarfaty, I., Benchmark for multi-cellular segmentation of bright field microscopy images. *BMC Bioinformatics* **2013,** *14*, 319.
214. Dimopoulos, S.; Mayer, C. E.; Rudolf, F.; Stelling, J., Accurate cell segmentation in microscopy images using membrane patterns. *Bioinformatics* **2014,** *30* (18), 2644-51.
215. Xing, F.; Yang, L., Robust Nucleus/Cell Detection and Segmentation in Digital Pathology and Microscopy Images: A Comprehensive Review. *IEEE Rev Biomed Eng* **2016,** *9*, 234-63.

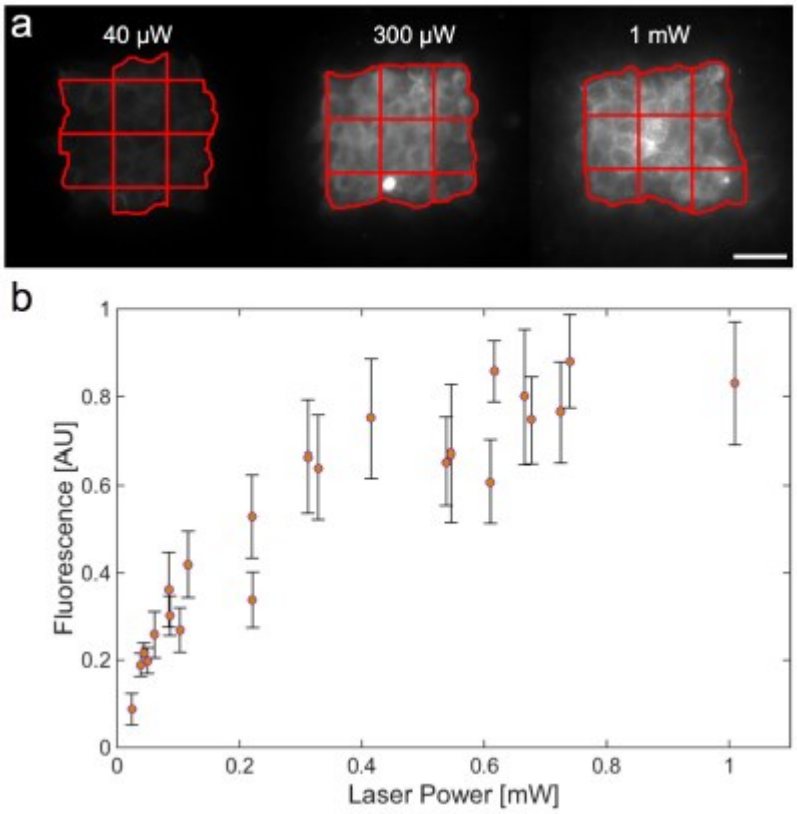
Annexe 1 : Informations supplémentaires publiées avec l'article *Live single cell laser tag*



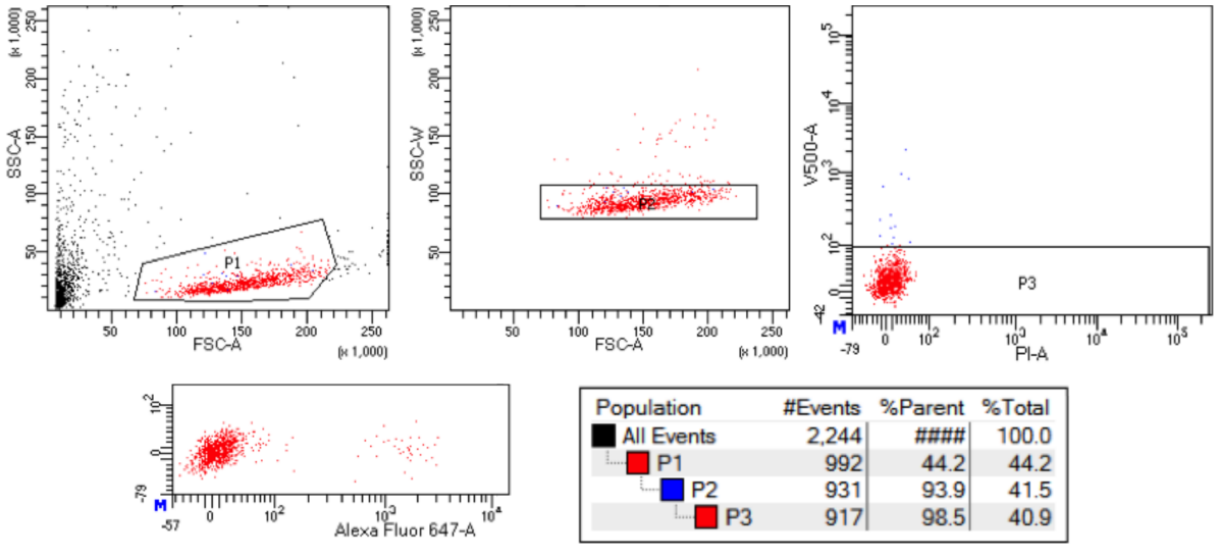
Supplementary Figure 1: Electron dense staining of isolated cells for electron microscopy imaging. CLaP was used to label cells for electron microscopy. (a) Control, non-tagged cells. (b) Positively stained cells with streptavidin-HRP are surrounded by a number of dark DAB precipitates, close to their filopodia where we expect the streptavidin HRP to be bound. (c) In an independent experiment, we used silver-enhanced 6 nm immunoGold-Streptavidin to reveal cells. Scale bars: 2µm.



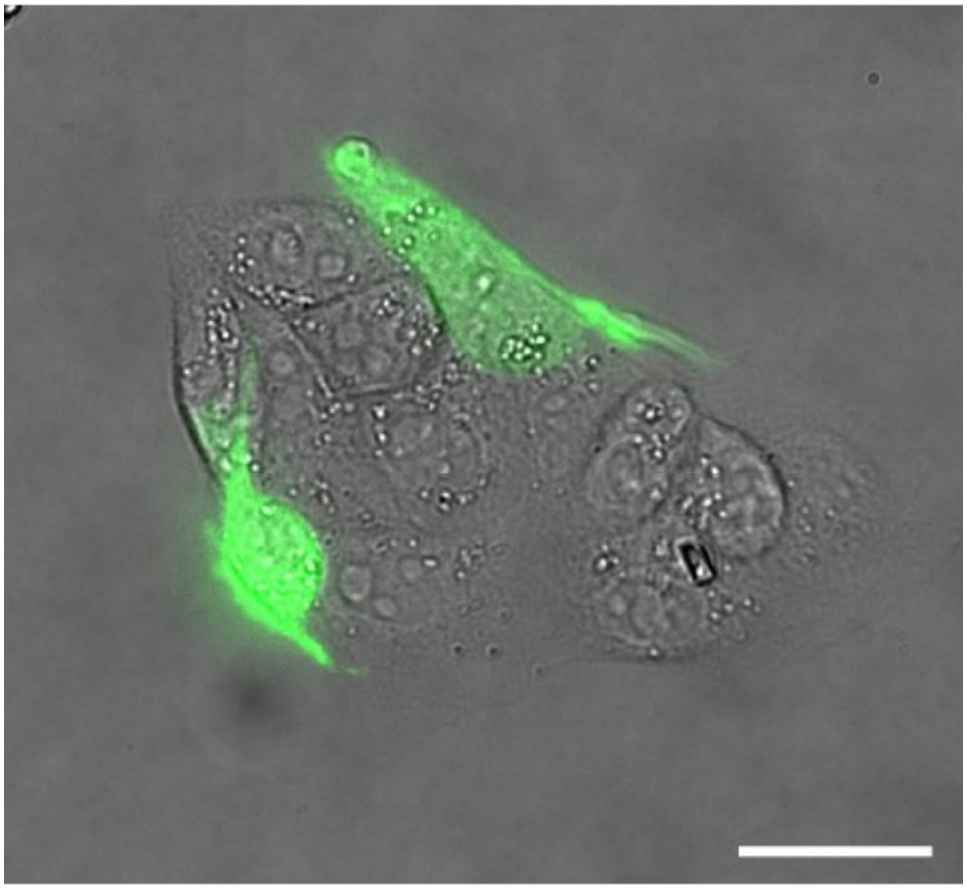
Supplementary Figure 2: CLaP tags are resistant to routine cell culture procedures. In order to test whether biotin tags withstand routine cell culture protocols, we illuminated individual ARPE-19 cells to crosslink biotin on their membrane. We then incubated cells in trypsin to detach them from the substrate. Cells were then plated in a new dish and let overnight in an incubator before adding Streptavidin-Alexa-647 to the medium and imaging on the following day. Isolated labeled cells (magenta) were readily identified among the large population of untagged cells stained with WGA (green). Scale bar, 20 µm.



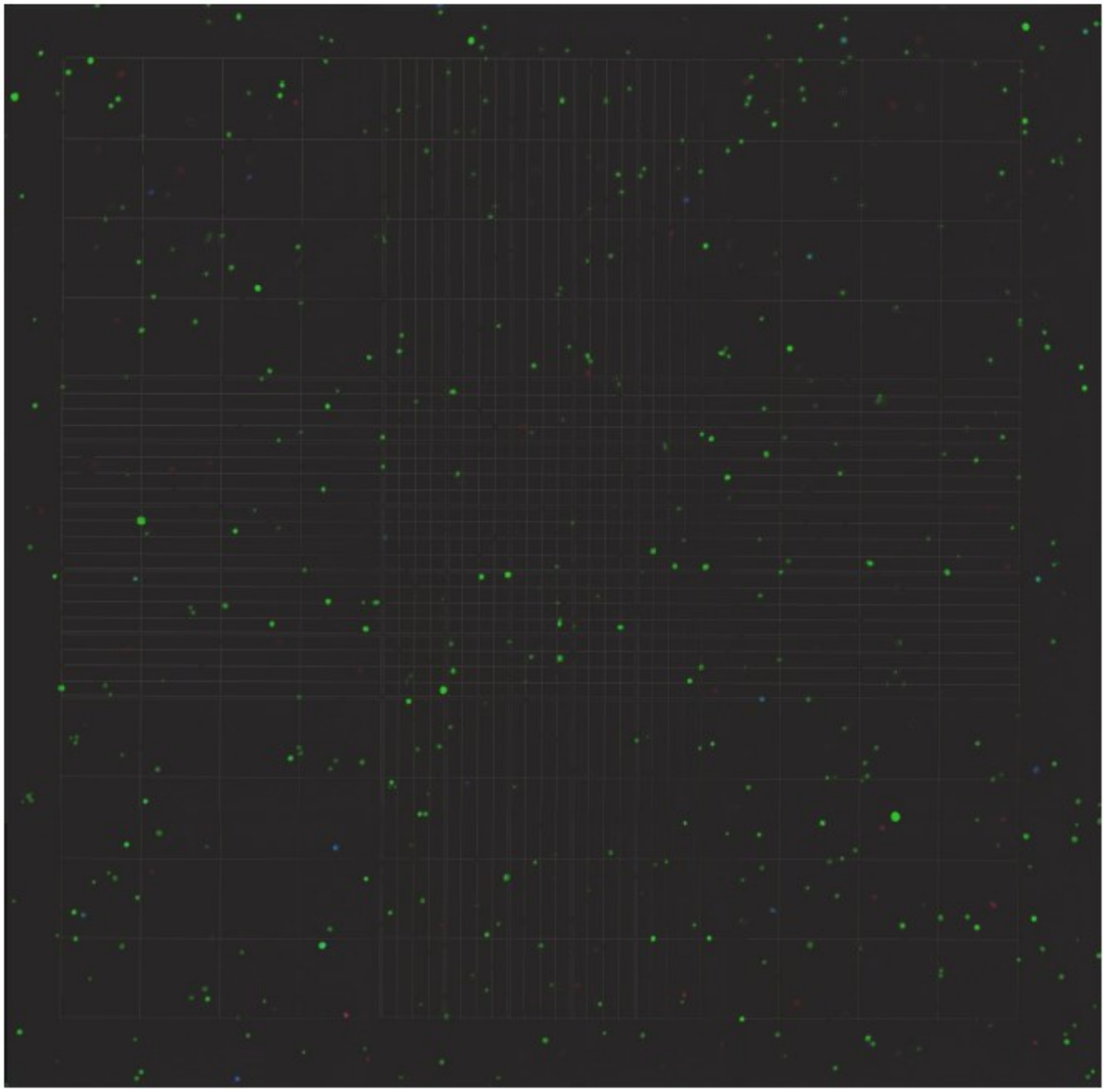
Supplementary Figure 3: Laser power calibration. (a) To calibrate the fluorescence intensity obtained with CLaP, we tagged square regions (100×100 µm$_2$) of MDCK cells using a 0.4 NA objective, scanning the sample at 170 µm/s, and using different laser intensities in each square: 40 µW, 300 µW and 1 mW. Laser intensity was measured at the objective focal plane and monitored at 10 Hz in order to control for fluctuations. After washing and incubating with Alexa-647-Streptavidin, we imaged each square region using epifluorescence with a 40XNA0.95 objective and 10s exposure time for all square regions. We chose objective magnification and spacing between squares to avoid photobleaching nearby squares during imaging. For similar reasons, the microscope field of view was placed around each square using prior knowledge of the pattern location, and focusing was done using brightfield illumination. Fluorescence intensity quantification of CLaP tagged cells was automated using MATLAB (MathWorks) to avoid subjective bias. A binary mask was created using the Otsu$_1$ algorithm followed by a morphological opening (using a disk-shaped structuring element of 5 µm radius), and a hole filling operation. Only the largest object in the binary image was kept. To assess fluorescence fluctuations within each pattern, the mask was subdivided in smaller square regions of 40×40 µm$_2$ as shown. Subregions smaller than 500µm$_2$ were discarded from the analysis. Scale bar, 40 µm. (b) Fluorescence within each pattern was characterized as a function of laser power. Markers and bars indicate the median and standard deviation of CLaP fluorescence of square regions.



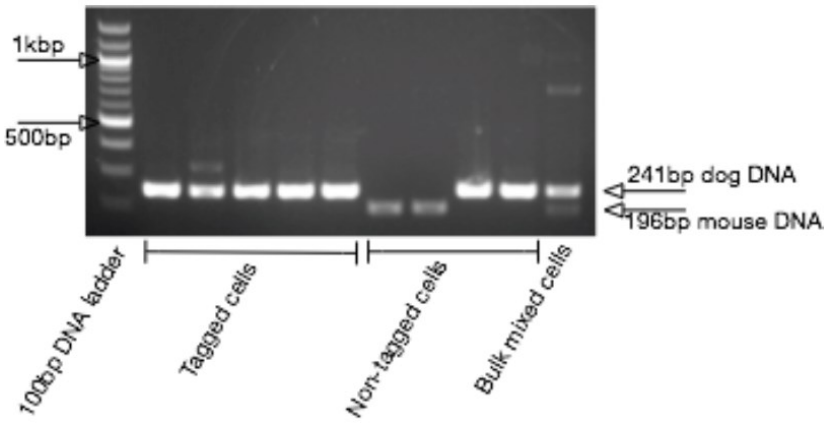
Supplementary Figure 4: Gatings for fluorescence activated cell sorting. Three standard gates were defined to count exclusively events originated from isolated viable cells. Gate P1 was built from the plot of side-scattered light (SSC) peak area vs. forward-scattered light (FSC) peak area, for discriminating whole cells from other particles and debris. A second gate (P2), used to discard doublets and other clusters, was defined using the graph of SSC peak width vs. FSC peak area. Finally, dead cells were filtered out with gate P3, which was defined from the peak area of the v500 channel (v500-A) which measured SYTOX Blue labeling.



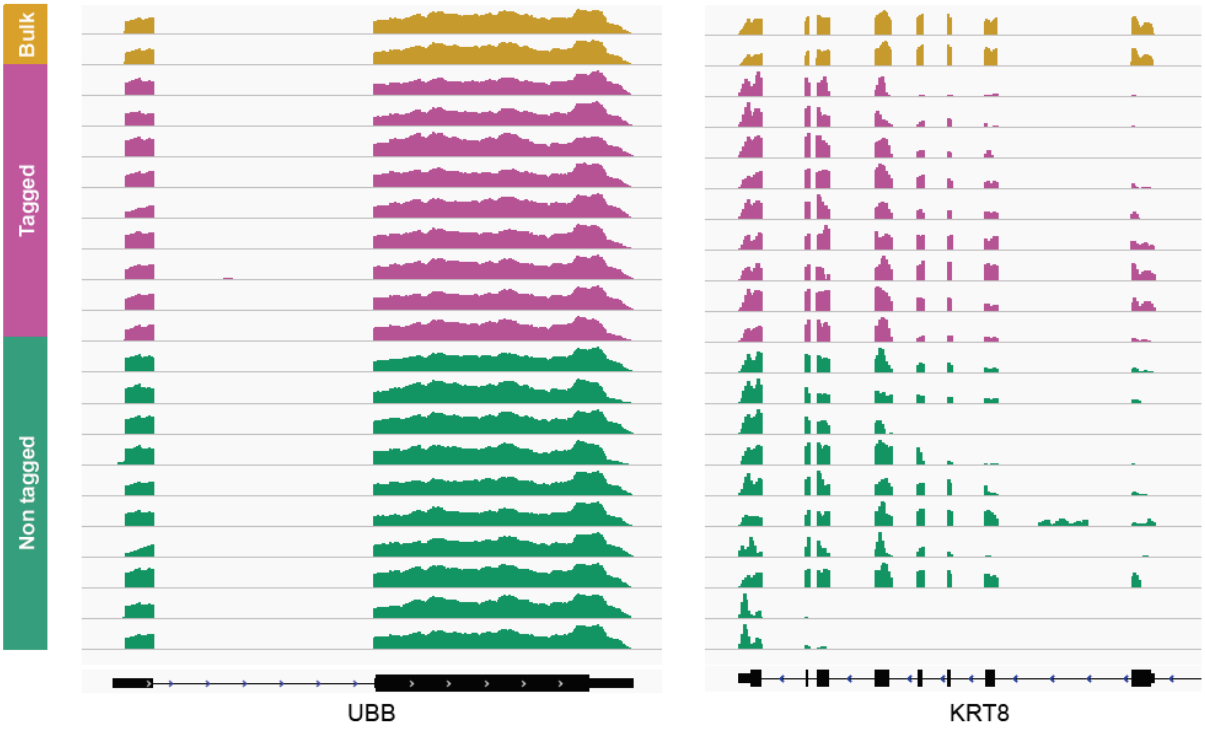
Supplementary Figure 5: mixed cell culture for single-cell microfluidics and capture. 3T3 fibroblasts expressing mNeonGreen (green), and non-fluorescent MDCK cells were co-cultured. Image illustrates mixed cell clusters with cell types that can be visually distinguished. Only MDCK cells (non-fluorescent in the image) were targeted by CLaP in subsequent experiments. Scale bar 20µm.



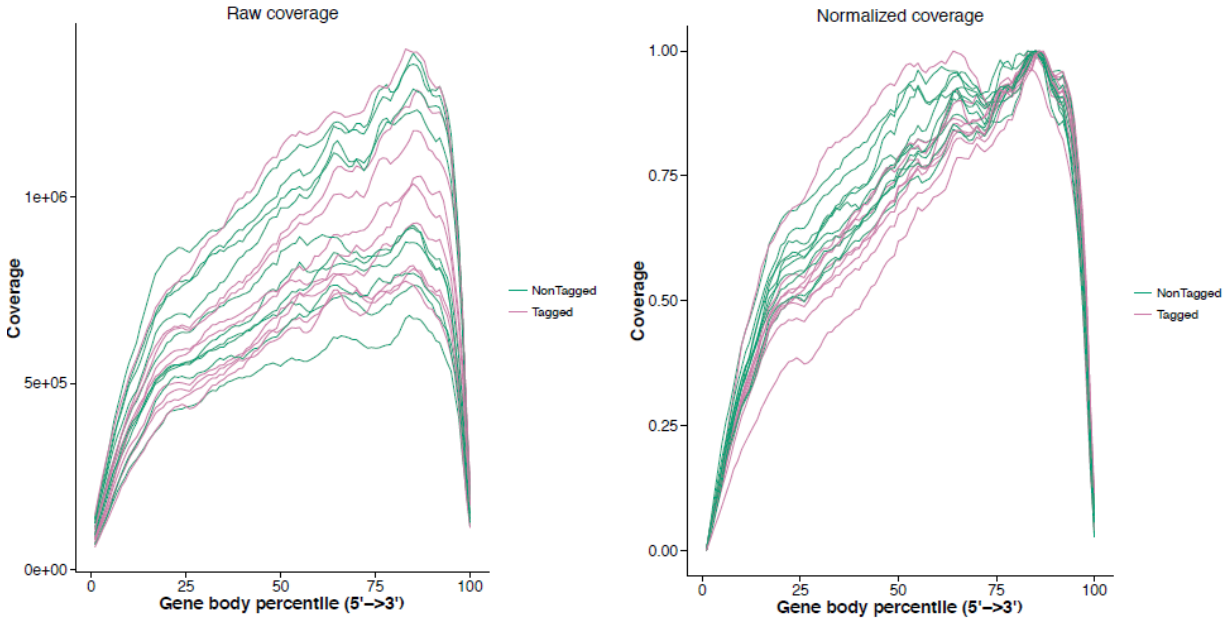
Supplementary Figure 6: Viability assay prior to C1 chip sorting. After CLaP, mixed cells were suspended using EDTA 10 mM and loaded in a C1 Fluidigm chip to isolate CLaP tagged cells from the rest. Live cells appear in green, dead cells appear in red, and alexa-647 streptavidin positive cells appear in blue.



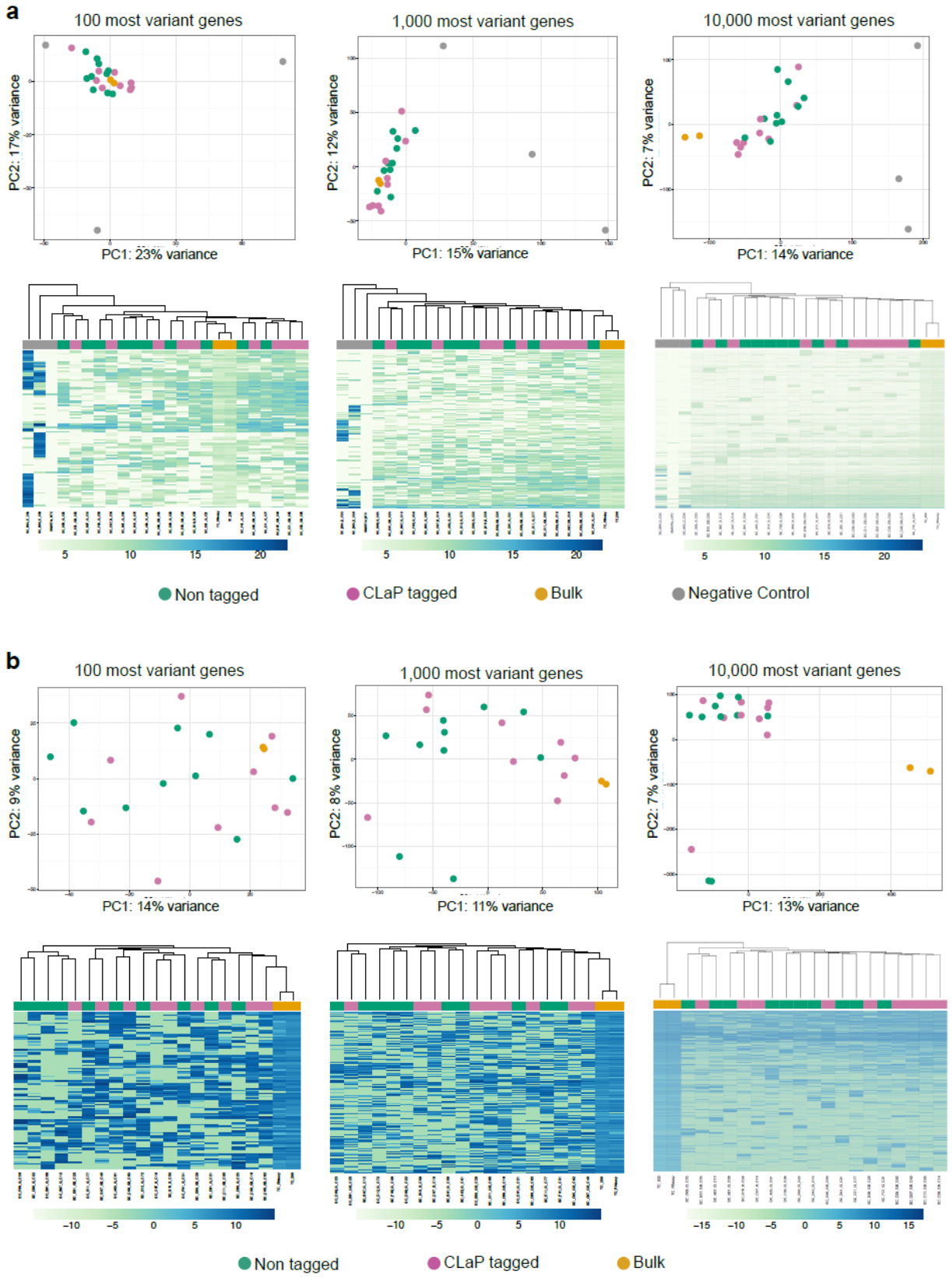
Supplementary Figure 7: Identification of species of origin in co-culture experiment. Cells from different species were co-cultured, tagged, isolated and analyzed by PCR. Here, the complete gel corresponding to figure 3c in the main text is shown, including molecular markers.



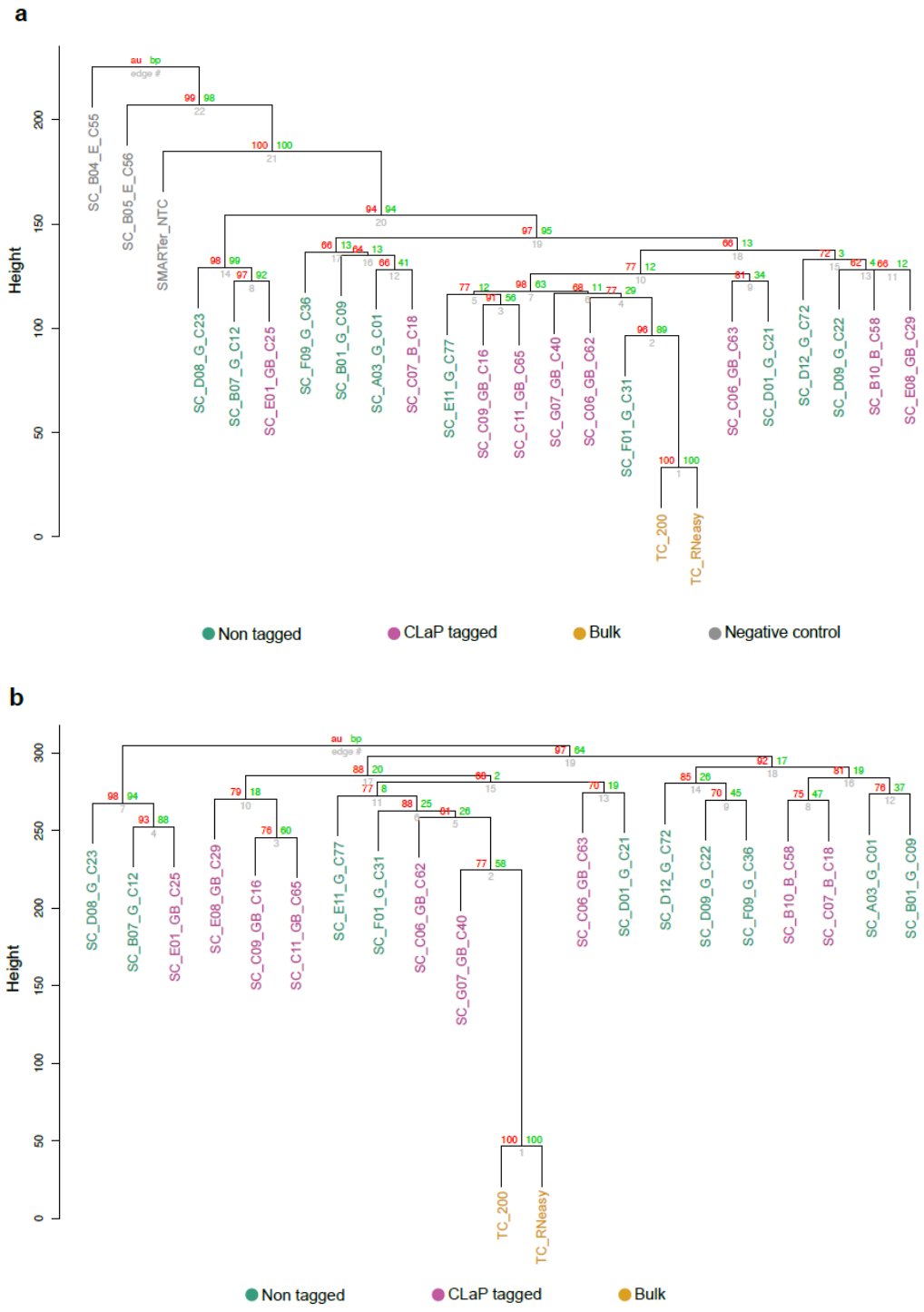
Supplementary Figure 8: Single-cell transcriptomic analysis. RNA-Seq data for one highly expressed gene (*UBB*) and one RPE marker (*KRT8*) from bulk (yellow), tagged single-cells (magenta) and non-tagged single cells (green).



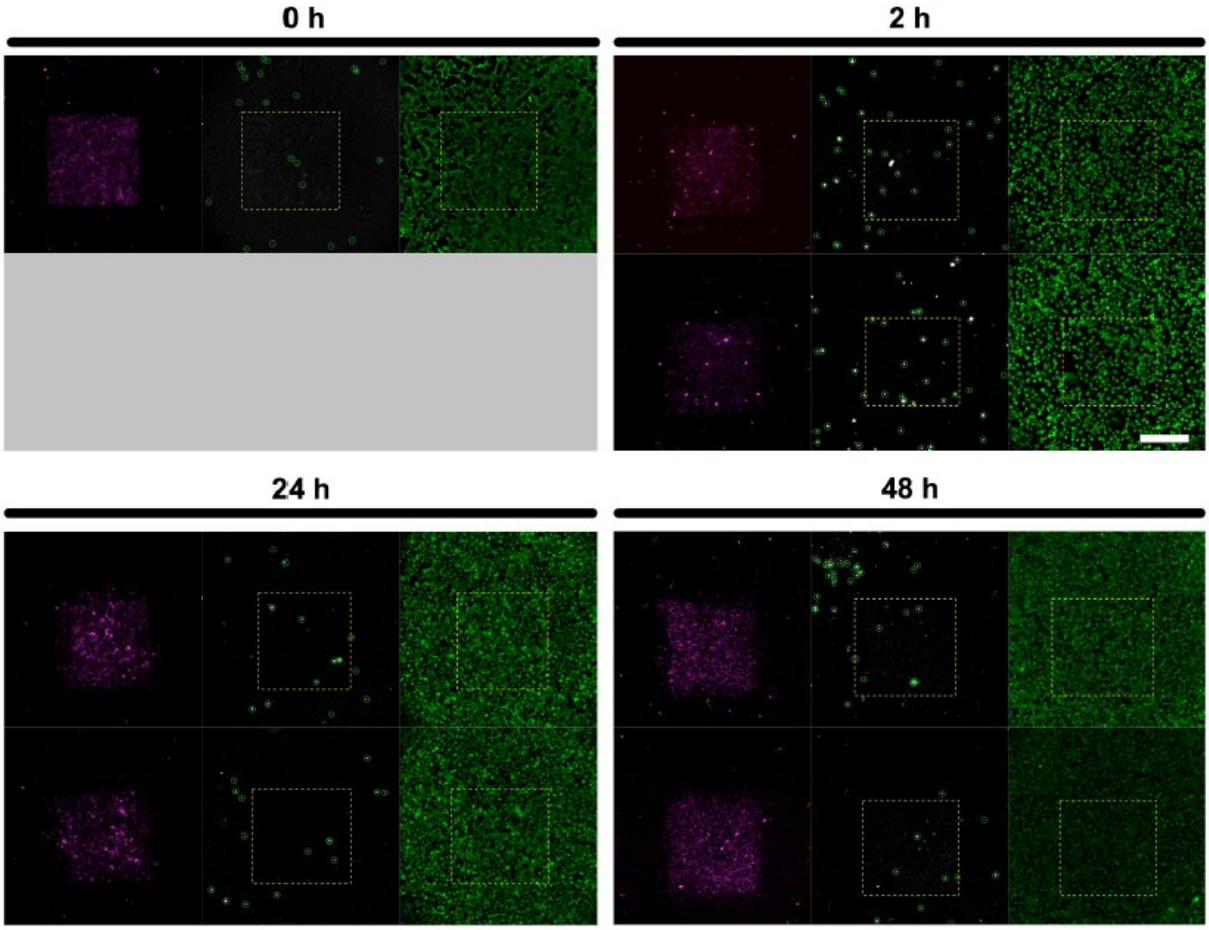
Supplementary Figure 9: Raw and normalized coverage uniformity over gene body. Using RSeQC (Bioinformatics 28 (16:2184-5), 2012), all transcripts were scaled to 100 nucleotides and the number of reads covering each nucleotide position was computed. The slight 3' bias observed is expected, since it has previously been reported for SMARTer Ultra Low RNA kit.


a
100 most variant genes
1,000 most variant genes
10,000 most variant genes
PC2: 17% variance
PC1: 23% variance
PC2: 12% variance
PC1: 15% variance
PC2: 7% variance
PC1: 14% variance
5
10
15
20
Non tagged
CLaP tagged
Bulk
Negative Control
b
100 most variant genes
1,000 most variant genes
10,000 most variant genes
PC2: 9% variance
PC1: 14% variance
PC2: 8% variance
PC1: 11% variance
PC2: 7% variance
PC1: 13% variance
−10
−5
0
5
10
−15
Non tagged
CLaP tagged
Bulk

Supplementary Figure 10: Effects of CLaP on gene expression. Global effects of CLaP on cells were evaluated by unsupervised clustering of samples based on expression profiles, usingvariable number of genes (from 100 to 10,000). Clustering was performed either including (a) or excluding (b) negative controls. Tagged and untagged cells are consistently clustered together, indicating no major expression changes associated with the procedure. Bulk: cDNA synthesized from 5 ng of purified total RNA derived from cells before capture in the C1 system, and a bulk cell control derived from approximately 20 lysed cells also before capture in the C1 system. Negative controls: libraries were prepared from cDNA synthesized from samples corresponding to two empty (no cell captured) positions in the C1 chip, producing low yield of cDNA (<0.5ng), as well as an ERCC spike only cDNA synthesis control.



Supplementary Figure 11: Multiscale bootstrapping of gene expression clustering, performed with the R package pvclust (Bioinformatics 22(12:1540-2), 2006). In red, the approximately unbiased (AU) p-value is represented. Bootstrapping was performed using the 1,000 most variant genes, including (a) or excluding (b) negative control samples.

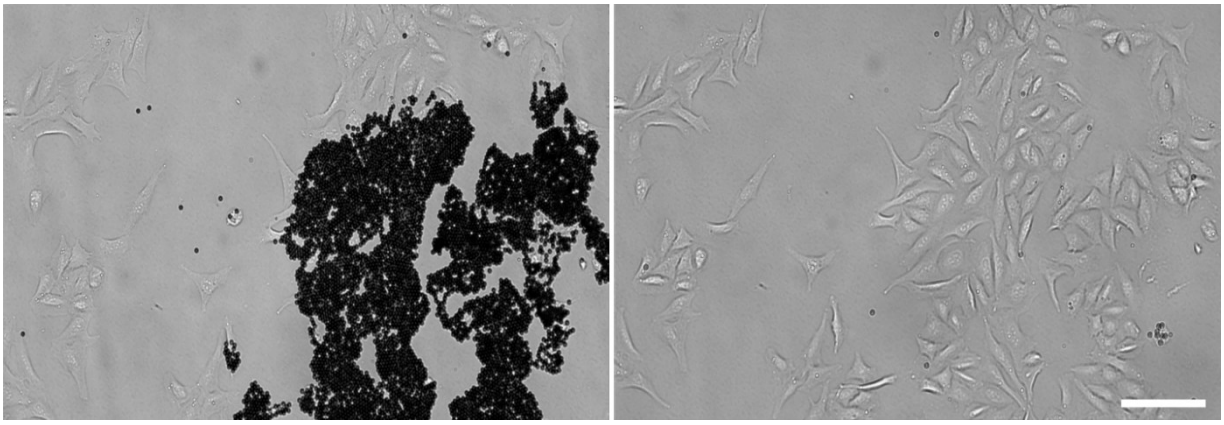


Supplementary Figure 12: Complementary viability analysis. As a complement for the viability study shown in Fig. 2c in the main text, we provide here the complete series of images used for quantification. At each time point, two samples were analyzed, except for 0h. For each sample we show three images (from left to right): Streptavidin-Alexa647 expression, PI expression and detections by the cell-counting algorithm, and Calcein AM expression. The rectangular regions displayed with dotted yellow lines represent the region subjected to CLaP, automatically segmented from the streptavidin-ALexa647 channel. Scale bar: 300µm.

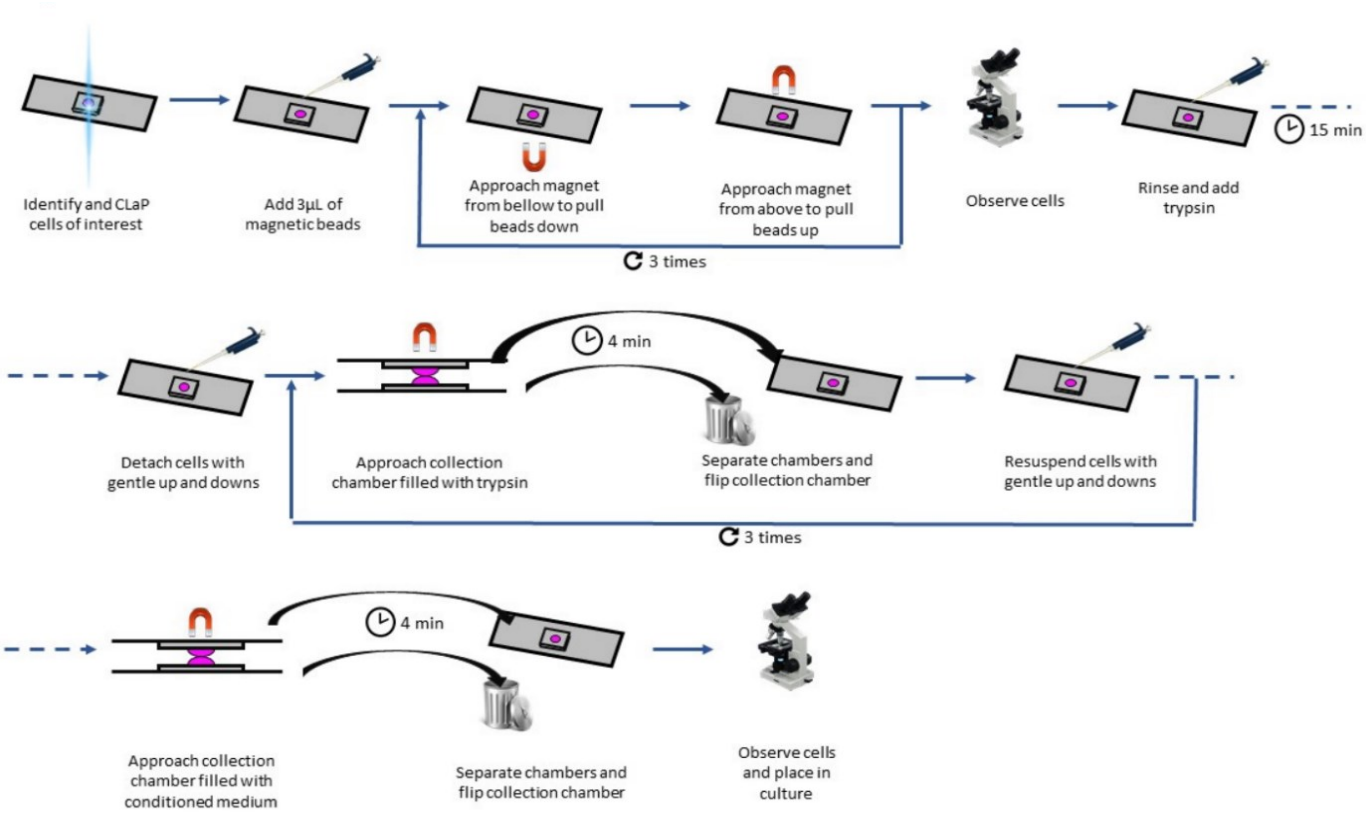
Supplementary Note 1: Photobleaching based crosslinking

As opposed to other studies that require the formation of nitrene groups upon UV illumination, crosslinking in CLaP is based on photobleaching fluorescein to create free radicals. The process that leads to the creation of free radicals using fluorescein has been extensively described by Song et al. Briefly, upon excitation by light absorption, molecules reach a singlet-excited state, which decays emitting fluorescence. These singlet-excited molecules can cross to an unstable triplet-excited state, causing photobleaching. When two triplet-sate molecules react, they produce semi oxidized and semi reduced radical forms of fluorescein. Both of these molecules have free electrons and are extremely unstable, reacting with molecules in their close vicinity to form stable compounds. Fluorescein is of particular interest for this specific use as it has a very low resistance to photobleaching and requires low illumination power to obtain free radicals. Biotin-4-fluorescein molecule has a short spacer arm between fluorescein and biotin, which allows the biotin to fit the binding pocket of streptavidin despite the presence the fluorescein tag.

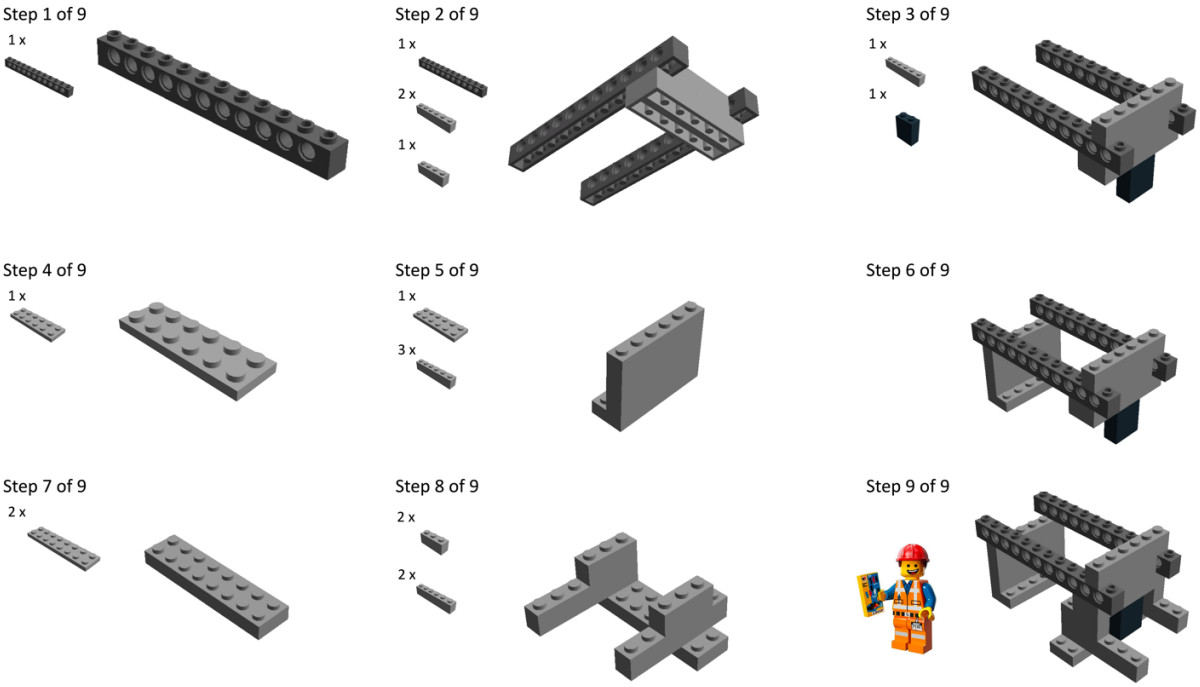
Annexe 2 : Informations supplémentaires publiées avec l'article *Opto-magnetic capture of individual cells based on visual phenotypes*



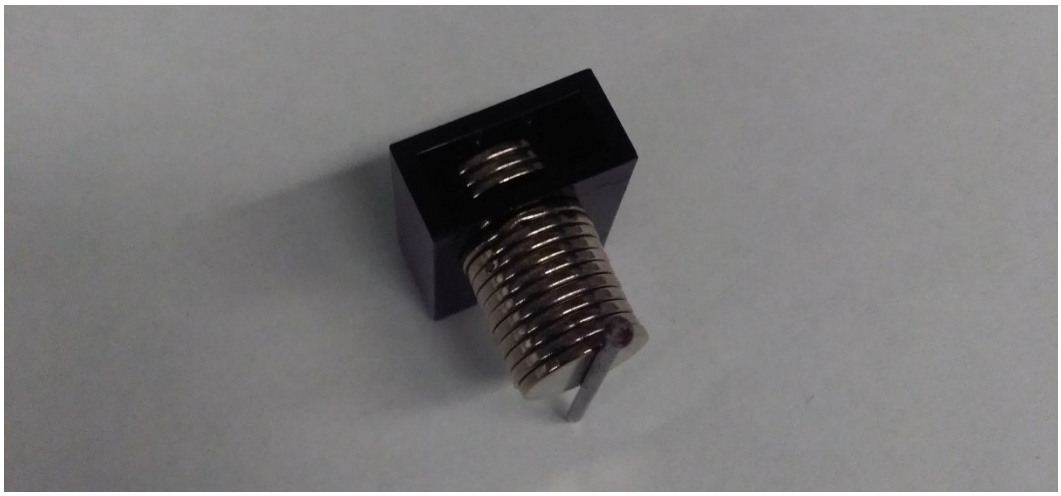
Supplementary Figure 13. Left: A fraction of cells in a culture was tagged with cleavable magnetic beads using scMOCa. Right: after incubation with DNAse, beads are cleaved and cells remain free of beads. Scale bar: 200µm.



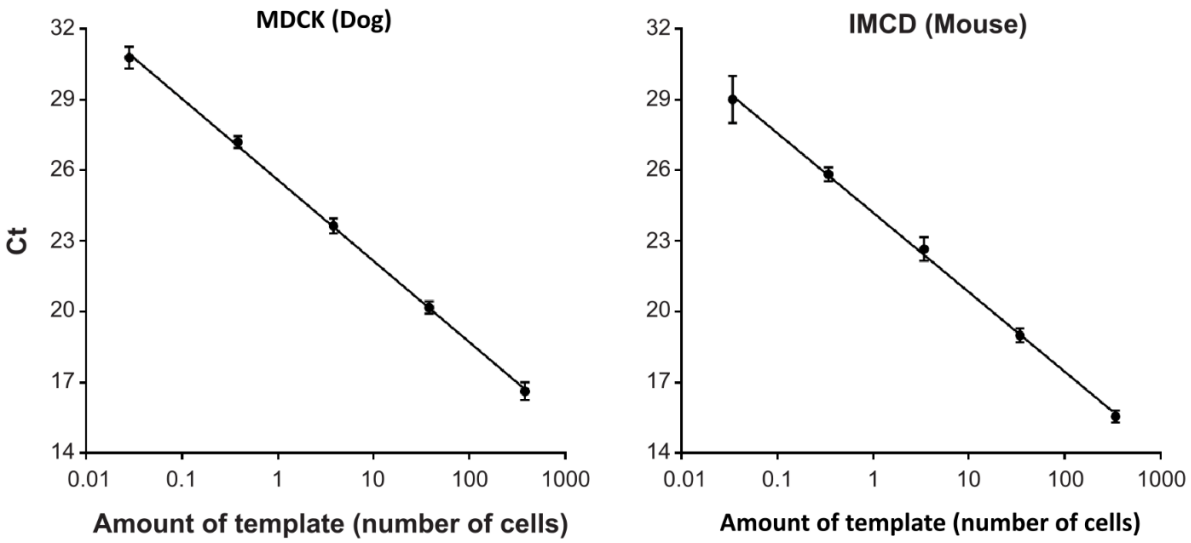
Supplementary figure 14. Step by step protocol to tag and isolate cells using scMOCa.



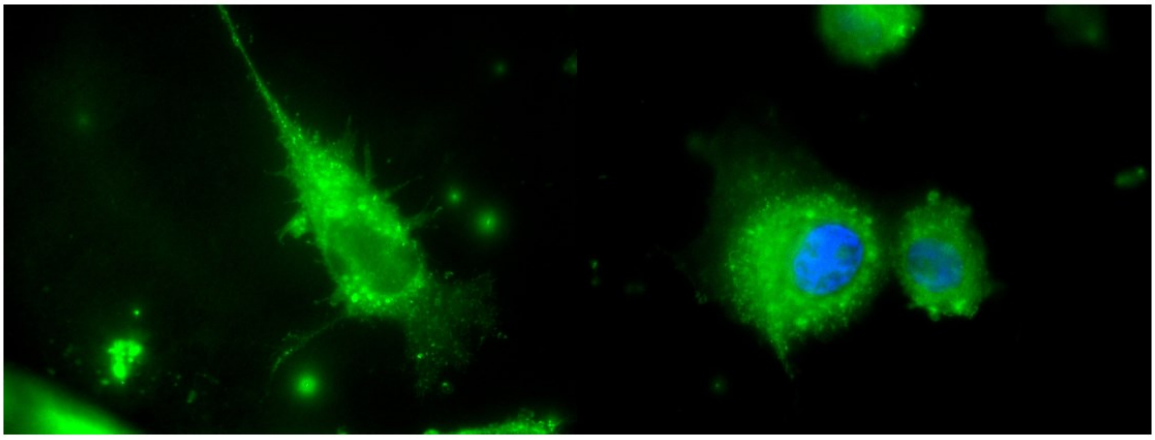
Supplementary figure 15. Instructions to create a simple platform to hold both chambers, the magnets and the nail.



Supplementary figure 16. Three magnets were inserted inside a hollow Lego brick to magnetically hold the rest of the pile in position for sorting.



Supplementary figure 17. Calibration curves used to calculate the number of cells from qPCR product. Abscises values are in the range of 1/100th of a cell to 100 cells which highlight that our experimental conditions allow detection of an amount of DNA corresponding to less than a cell.



Supplementary figure 18. Two examples of cells that were killed with sodium azide then stained with mitotracker green and Hoechst. This negative control confirms the cells shown in figure 6 are live cells. When cells are alive, mitotracker green stains filament shaped mitochondria, whereas when cells are dead, the spotty signal it generates is distributed everywhere within the cytoplasm.