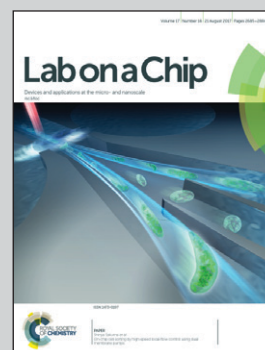


Featuring work from the Lab for Bio-integrated Photonics,
Dr. Matjaž Humar, Jožef Stefan Institute, Slovenia.

Spectral reading of optical resonance-encoded cells in microfluidics

We demonstrate the feasibility of large-scale cell barcoding and identification using intracellular micro-resonators, which produce unique spectral features or barcodes. A microfluidic chip was used to first measure the fluorescence resonance peaks emitted from cells and later to identify the individual cells.

As featured in:



See Seok Hyun Yun *et al.*,
Lab Chip, 2017, 17, 2777.


 Cite this: *Lab Chip*, 2017, 17, 2777

Spectral reading of optical resonance-encoded cells in microfluidics†

 Matjaž Humar, ^{‡abc} Avinash Upadhyay^{‡ad} and Seok Hyun Yun^{*ae}

The ability to label individual cells is useful for single-cell-level studies of complex cellular interactions and heterogeneity. Optically readable cell labeling is attractive as it can be investigated non-invasively and repeatedly at high speeds. Here, we demonstrate the feasibility of large-scale cell barcoding and identification using fluorescent polystyrene microbeads loaded into cells. Intracellular beads with different diameters in a range of 5 to 12 μm generate spectrally distinguished features or barcodes. A microfluidic chip was used to measure fluorescence resonance peaks emitted from individual cells. An algorithm comparing the peak wavelengths to a reference barcode library allowed barcode identification with high accuracy. This work provides a guideline to increase the number of unique identifiers and reduce various false-positive and false-negative errors.

 Received 2nd March 2017,
Accepted 28th June 2017

DOI: 10.1039/c7lc00220c

rsc.li/loc

Introduction

Barcoding in biomedical sciences is a labeling technique to identify objects from otherwise non-distinguishable population. DNA barcoding was developed for the identification of species and tracing the migratory behaviors of animals, such as birds, but its applications have been extended to biological cells for tracing their lineage, genetic mutations, and migration.^{1,2} Since DNA barcodes are encoded in the genome, their readout is typically performed by DNA sequencing, which requires cell lysis. Possibly, DNA barcodes can be detected by using fluorescence *in situ* hybridization,³ but this method requires multiple stages of staining and washing, making measurements slow and not readily applicable to cells in living organisms. In mass cytometry, rare-earth metals with distinct atomic weights are widely used for labeling target molecules and cells.^{4,5} Multiplexing up to 96 labels has been demonstrated,⁶ and the number of identifiable labels may be further increased.

In contrast to the DNA and rare-earth metal barcodes, optically accessible barcodes allow non-invasive, rapid readouts and offer the possibility of real-time and time-lapse analyses. Among the variety of optical barcoding schemes developed to date,⁷ graphical encoding relies on micro- or nano-patterns that are distinguished by using optical microscopy or nanoscopy.^{8–10} However, this method is prone to error unless the barcode patterns are clearly visualized and thus, may not be applicable to *in vivo* applications. Information can also be encoded by the magnitudes of forward and side scattering of specifically engineered layered spheres¹¹ and luminescence lifetimes.¹² Spectrometric barcoding is based on the combination of non- or minimally-overlapping emission spectra of different organic dyes, quantum dots,^{13,14} Raman scattering molecules,¹⁵ and lanthanides.¹⁶ Stochastic expression of multiple fluorescent proteins^{17,18} can produce numerous colors distinguished by ratiometric spectral analysis and can be combined with spatial mapping.¹⁹ Fluorescent beads made using three dyes and ten different concentration levels to generate 500 possible combinations have been commercially developed.²⁰ While such beads can in principle be used for cellular barcoding, intensity-graded coding is less reliable than spectroscopic barcoding especially for intra-tissue and *in vivo* settings, where optical scattering and absorption in tissue and photo-bleaching of the incorporated dyes may diminish the accuracy of intensity-based readouts.

Among various optical barcoding schemes, coupling of fluorescence to an optical microcavity is an attractive approach because it can generate narrowband spectral features defined by cavity resonance. In particular, fluorescent microbeads can support whispering-gallery modes (WGMs) and generate feature-rich resonance spectra²¹ with good

^a Wellman Center for Photomedicine, Harvard Medical School, Massachusetts General Hospital, 65 Landsdowne St. UP-5, Cambridge, Massachusetts 02139, USA. E-mail: syun@hms.harvard.edu

^b Condensed Matter Department, J. Stefan Institute, Jamova 39, SI-1000 Ljubljana, Slovenia

^c Faculty of Mathematics and Physics, University of Ljubljana, Jadranska 19, SI-1000, Ljubljana, Slovenia

^d Research School of Engineering, Australian National University, North Road Canberra, Canberra, Australian Capital Territory 0200, Australia

^e Harvard-MIT Health Sciences and Technology, 77 Massachusetts Avenue Cambridge, Cambridge, Massachusetts 02139, USA

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c7lc00220c

‡ These authors contributed equally.

photostability, making them suitable for cellular barcoding. Recently, we and others have shown that dye-containing microspheres can be incorporated into cells and generate stable WGM spectra in both spontaneous (fluorescence) and stimulated emission (lasing) regimes.^{22–24} In the event of cell division, beads can be passed down to daughter cell(s) over several generations with very small spectral shifts (<30 pm) of their WGM peaks.²⁵ Here, we report the feasibility of reading WGM-based spectral barcodes from cells in a microfluidic channel. Our experiments revealed several sources of readout errors and provided insights into how and to what extent the errors may be reduced for large-scale cellular barcoding and high-throughput readouts with improved accuracy.

Results

Loading microspheres into cells

A commercial batch of fluorescent polystyrene beads (Thermo Scientific, Fluoro-Max, with a wide range of sizes from 5 to 12 μm , the coefficient of variation of the particle diameter, CV = 18%) was used for each labeling experiment. The bead sizes measured from their fluorescence microscopy images exhibit a Gaussian distribution with a mean diameter of 8.79 μm and a full-width-half-maximum of 2.92 μm (standard deviation of 1.24 μm) (Fig. 1a), consistent with the manufacturer's specification. After washing with ethanol and water, the microbeads were immersed in 1% w/v aqueous solution of poly-L-lysine hydrobromide (MW 30 000–70 000, Sigma) for 30 min to form a poly-L-lysine coating on the surface. They were washed three times with water, centrifuged at 5000g for 5 min, and transferred to phosphate buffered saline (PBS) solution. HeLa cells were grown at 37 °C with 5% CO₂ in full growth medium (DMEM medium with 10% fetal bovine serum and 1% pen-strep). The microbeads in PBS (1.84×10^6 beads per ml) were added to the cell culture medium (8.57×10^4 cells per ml) and incubated for 24 h, by which the cells reached ~70% confluency. After the 24 h incubation, we typically found that about 80–85% HeLa cells contained at least one bead, and ~50% of the bead-containing cells had more than one bead (Fig. 1b). The measured distribution of the number of beads is

fitted well with a Poisson distribution with an average of 1.7 beads per cell (Fig. 1c). This indicates that the internalization of beads can be described as a random process, which is independent of the bead size²⁴ and unaffected by the presence of other beads inside a cell. The cells were washed with PBS to remove non-internalized beads, detached with trypsin, centrifuged at 125g for 5 min, and re-dispersed in PBS medium containing 1% w/v bovine serum albumin (BSA). BSA reduces sticking of cells to the internal surfaces of syringes, tubing and microchannel walls. The dispersion was filtered through a 35 μm mesh to remove clusters of cells and other debris before loading into 100 μl glass syringes (Gastight 1710).

Spectral readout of resonance-encoded cells

The microfluidic chip used was made of a cyclic olefin copolymer (ChipShop) and had 58.5 mm-long channels on a 140 μm -thick bottom plate (Fig. 2a). Each channel had a square cross-section of 50 μm by 50 μm . A syringe filled with a suspension of cells (2.1×10^5 cells per ml) was connected to one end of the channel *via* 8 cm-long tubing with an inner diameter of 1.0 mm and olive connectors. Another syringe filled with PBS was connected to the other end of the channel using the same type of tubing. The syringes were placed on two separate syringe pumps (KDS Legato 100). Air bubbles inside the syringes, tubes, and channel were carefully removed. The chip was mounted on a custom-built inverted microscope setup (Fig. 2b). For fluorescence excitation, the output of a continuous-wave 491 nm laser (Cobolt Dual Calypso) was delivered to the middle of the channel through an objective lens (10 \times , 0.25 NA, Nikon). The beam was slightly defocused so that the beam size at the channel would be 50 μm , which matched the width of the channel. The fluorescence emission from the channel at the focus of the objective lens was directed to a CCD camera (Mightex, MT9001) for bright-field and fluorescence imaging (Fig. 2c) and to a diffraction-grating spectrometer (Andor Shamrock 303i-B) with a spectral resolution of 0.05 nm. The fluorescence spectra were recorded at a rate of 28 spectra per s with an exposure time of 36 ms per measurement.

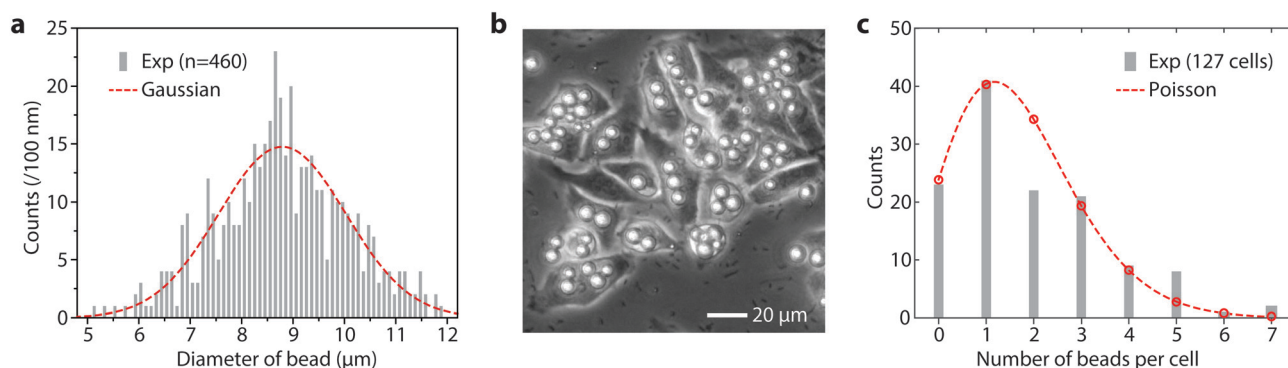


Fig. 1 (a) Histogram distribution of the bead diameter measured from fluorescence microscopy images of 460 beads. Bin size: 100 nm; dashed curve (red): Gaussian distribution fit. (b) HeLa cells grown with polystyrene beads in a dish. (c) Histogram of the number of beads per cell. Dashed curve (red): best-fit Poisson distribution (mean = 1.7).

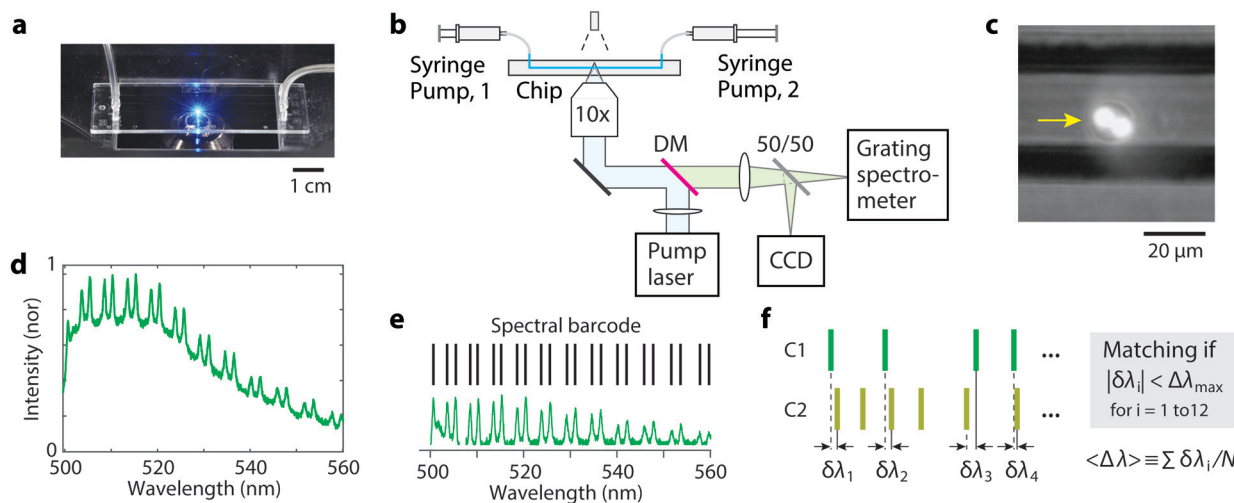


Fig. 2 (a) The microfluidic chip mounted on a microscope. (b) Schematic of the optical setup. Pump laser: a continuous-wave laser at 491 nm. DM: dichroic mirror. (c) CCD image of a cell (arrow) with two fluorescent beads, superimposed on a bright-field image of the channel. (d) A typical fluorescence spectrum from a cell containing a single bead, featuring a periodic double-peak WGM structure (TM and TM polarizations) on the top of a broad fluorescence background. (e) The group of WGM peak wavelengths identified from a fluorescence-subtracted spectrum constituting the “spectral barcode” of the cell. (f) Two barcodes, C1 and C2, which are compared in terms of spectral differences, $\delta\lambda$, and deemed identical if $|\delta\lambda| < \Delta\lambda_{\max}$ for at least 12 resonance peaks. The average of $\Delta\lambda$ is denoted $\langle\Delta\lambda\rangle$.

There is more than one method to convert the spectrum to a unique identifier of the cell. In principle, the WGM spectrum can provide sufficient information to determine the bead diameter with sub-nanometer resolution by curve fitting based on the WGM theory.²² This method can distinguish cells based on the diameters of the beads. This fitting-based method was shown to be effective for individual single beads²² but can be unreliable for spectra originating from multiple beads. Another approach is to use the spectral curve itself as an identifier and to perform some sort of cross-correlation integral to determine the identity between two spectra. This method is also effective for single beads but far less robust when cells have multiple beads because the overall spectral shape is sensitive to fluorescence intensity and the relative intensities of individual beads within a cell can vary as the relative position of the two beads is changed, particularly with respect to the pump beam axis.

To enhance the robustness of tagging, we devised a simple, yet effective, protocol, described as follows. The spectrum of a cell containing one or a few beads shows many WGM resonance peaks on the top of a broad fluorescence spectrum over a spectral range of 500 to 580 nm (Fig. 2d). In post processing, the broadband background was subtracted from the measured spectrum, leaving only the WGM resonance peaks. The spectral peaks were fitted by Lorentzian line shape functions to locate the central positions of the peaks. The numerical array of the WGM peak wavelengths in an increasing order constitutes a “barcode” (Fig. 2e). We considered a WGM peak “clearly resolved” when the peak amplitude was at least 3-times larger than the root-mean-square (rms) value of the noise floor in the spectrum and the peak wavelength must be separated from its neighboring spectral peaks by at least 1 nm. Only the cells that exhibited 12 or

more clearly-resolved spectral peaks were considered “barcoded” successfully and included in further analysis. A typical barcoded cell contained 12 to 24 clearly-resolved spectral peaks or barcode lines.

We used the following protocol to determine whether two barcodes are identical or different. First, the number of barcode lines is compared. If the difference in the number of peaks is greater than 6, the two barcodes would be determined to be different or non-matching, because this situation would occur when the cells comprised either different number of beads or same number of beads but with very dissimilar sizes. On the other hand, when the difference is 6 or less, then their peak wavelengths are compared. For this comparison between the two barcodes, the one with the smaller number of spectral peaks (C1) is selected. For each reference peak of the first barcode, the closest peak in the second barcode (C2) is determined, and the spectral difference in wavelength, $\delta\lambda_i$, is calculated (Fig. 2f). The spectral difference is compared to a fixed parameter $\Delta\lambda_{\max}$, termed the *maximum spectral shift*. If $\delta\lambda_i$ is smaller than $\Delta\lambda_{\max}$ for at least 12 spectral peaks, the two barcodes would be regarded as identical or matching; if not, they would be deemed different. Thus, the criterion for barcode matching is $|\delta\lambda_i| < \Delta\lambda_{\max}$ for $i = 1$ to 12. We optimized the parameter $\Delta\lambda_{\max}$ experimentally to achieve the most desirable outcomes. The optimization step is described in more detail later. Since we compare only the central wavelength positions of the peaks, this algorithm is relatively insensitive to the relative intensities of spectral peaks.

Verification of barcode reading with stationary cells

As an initial test of the readout scheme, we loaded a suspension of cells into the microfluidic chip and sealed both ends of the

channel (Fig. 3a), so that the position of cells in the channel was not changed during measurements. The chip was mounted on a motorized stage, and the stage was translated in the forward and then backward directions at a speed of 0.9 mm s^{-1} while the fluorescence spectra were continuously recorded (Fig. 3b). Between the two scans, the order of cells was retained. In the forward run, a total of 87 cells containing beads passed through the laser-excitation zone. From the typical bead-uptake efficiency of 80–85% we measured for the HeLa cell line, we estimated that about 18 cells without any beads passed through the excitation zone during the measurement without being detected in the fluorescence measurement. Out of 87 detected cells (containing beads), 72 generated emission spectra, each with at least 12 clearly-resolved spectral peaks, satisfying the detection criterion defined earlier. Almost all single-bead and most two-bead cells satisfied the criterion. However, almost all cells containing 3 or more beads did not satisfy the detection criterion because of the overlap of spectral peaks and, hence, were rejected (not “detected”). The detected 72 cells constitute a reference group, among which ~ 50 cells had single beads and ~ 22 had two or more beads. In the backward run, 75 barcoded cells met the criterion and were “detected”. The reason why there were 3 more cells detected in the backward scan is presumably because of the small intensity fluctuations between the forward and backward measurements. These cells were not detected in the forward scan because they had less than 12 clearly-resolved spectral peaks, but some of the undetected peaks exceeded the detection threshold—3 times the rms inten-

sity noise—in the backward scan. Considering the 75 cells as a test sample group, we applied our barcode matching protocol described earlier with respect to the reference cells.

First, we analyzed the correlation between reference cells for various $\Delta\lambda_{\text{max}}$ values and found that $\Delta\lambda_{\text{max}} = \sim 100 \text{ pm}$ was suitable to resolve all the 72 cells without any correlation hit between two different cells (all cells had a unique barcode). A correlation map for $\Delta\lambda_{\text{max}} = 100 \text{ pm}$ was drawn (Fig. 3c), where the *x*- and *y*-axes represent cell indices in the temporal order in which they were measured, and a circular correlation hit mark is placed where the corresponding pair has identical barcodes according to the criterion for the given $\Delta\lambda_{\text{max}}$. Along the diagonal axis, each barcode is compared to its own self, so this axis is trivial and always filled with circular marks. On the other hand, any correlation hit in the non-diagonal space of the map is non-trivial and could result from two scenarios: (1) $\Delta\lambda_{\text{max}}$ is too large so that barcodes from different beads satisfy the criterion for matching, which leads to a false-positive error; or (2) the two cells have different beads but their diameters are too similar to be resolved with the given $\Delta\lambda_{\text{max}}$. Reducing $\Delta\lambda_{\text{max}}$ may decrease these false-positive counts, but this comes at the expense of increasing false-negative errors failing to identify the same cells. At $\Delta\lambda_{\text{max}} = 100 \text{ pm}$, the analysis for the backward-scan dataset yielded no false-positive error as clearly shown in a correlation map (Fig. 3d).

The same protocol was applied to analyze the cross-correlation between the sample and reference cells. 66 of 75

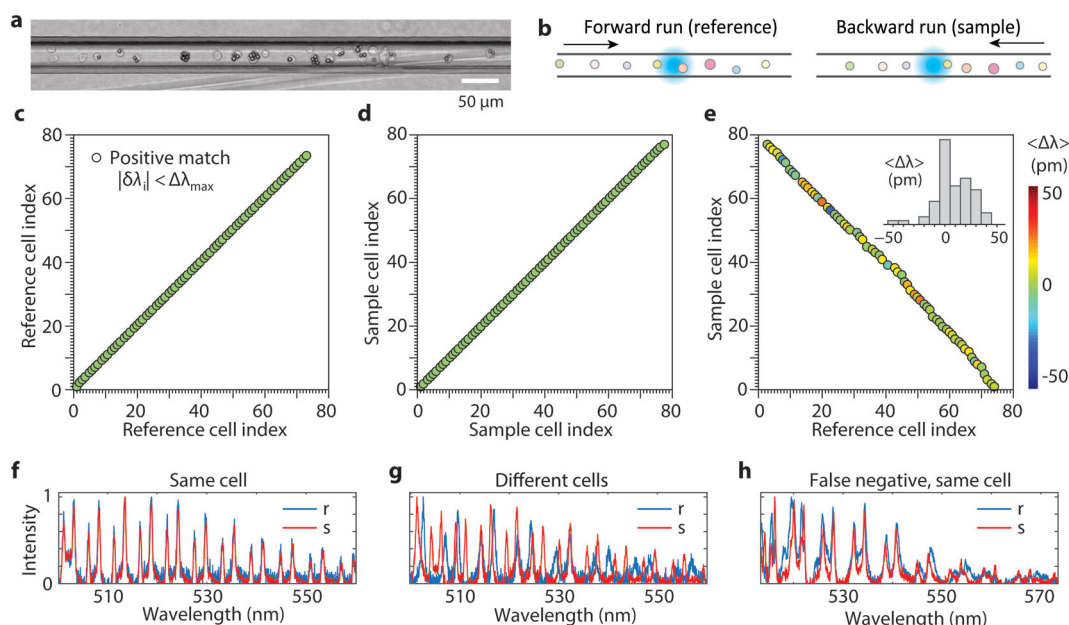


Fig. 3 (a) Micrograph of the channel loaded with cells. (b) Measurement strategy. Cells are scanned in the forward and then backward directions across the excitation beam (cyan). (c and d) Correlation maps obtained with $\Delta\lambda_{\text{max}} = 100 \text{ pm}$ for (c) the reference (forward flow) and (d) test sample (backward flow), and (e) cross correlation between reference and sample cells. Inset, histogram of the measured mean spectral shift of identified WGM peaks between reference and sample measurements of matching cells. Color represents the mean spectral shift. (f) An example set of spectra of a cell in reference (*r*) and sample (*s*) dataset, showing high cross-correlation between them. Fluorescence background was subtracted from all spectra. (g) Example of reference and sample spectra of two different cells with apparently low cross-correlation. (h) An exemplary failure mode (false-negative error) that originates from spectral shifts between the reference and sample spectra of a cell containing two beads.

cells in the test sample matched with different reference cells. This result is illustrated in a correlation map (Fig. 3e), where correlation marks appear along the negative-slope diagonal axis simply because the order of cells was reversed in the backward scan. There was no double-matching error; that is, each test cell matched to only one reference cell, and no two test cells matched to a single reference cell. As noted above, 3 in 75 cells did not have matching references, and additional 6 cells failed to find their matching cells (false-negative errors). To better understand the causes of the errors, we manually examined the measured spectra. Any pairs of matching cells exhibited well-overlapping spectra (Fig. 3f). The spectra of non-matching cells showed apparent differences (Fig. 3g). However, the 6 cells which failed to find their matching cells (false-negative errors) had their spectra slightly shifted between the reference and sample scans

(Fig. 3h). The magnitude of the shift was greater than 100 pm for some peaks, which reduced the number of matching peaks that satisfy $|\delta\lambda_i| < 100$ pm to less than 12. The spectral changes of >100 pm are presumably caused by the changes in the effective surrounding refractive index of the beads.

We calculated the mean spectral shift, $\langle\Delta\lambda\rangle = \sum\delta\lambda_i/N$, where the summation is over the entire set of spectral lines (N) in the barcode that satisfied $|\delta\lambda_i| < \Delta\lambda_{\max}$, where typically $N = 12$ to 24. The color of each circle in the cross-correlation map (Fig. 3e) represents the mean spectral shift of the test barcodes with respect to their matching reference barcodes. The histogram of the calculated values (inset of Fig. 3e) ranged from -50 to 50 pm and showed some indication of a red shift. Possible phenotype changes or temperature changes of the cells in the microfluidic channel may have caused the asymmetric shift, changing the surrounding refractive index of the beads.

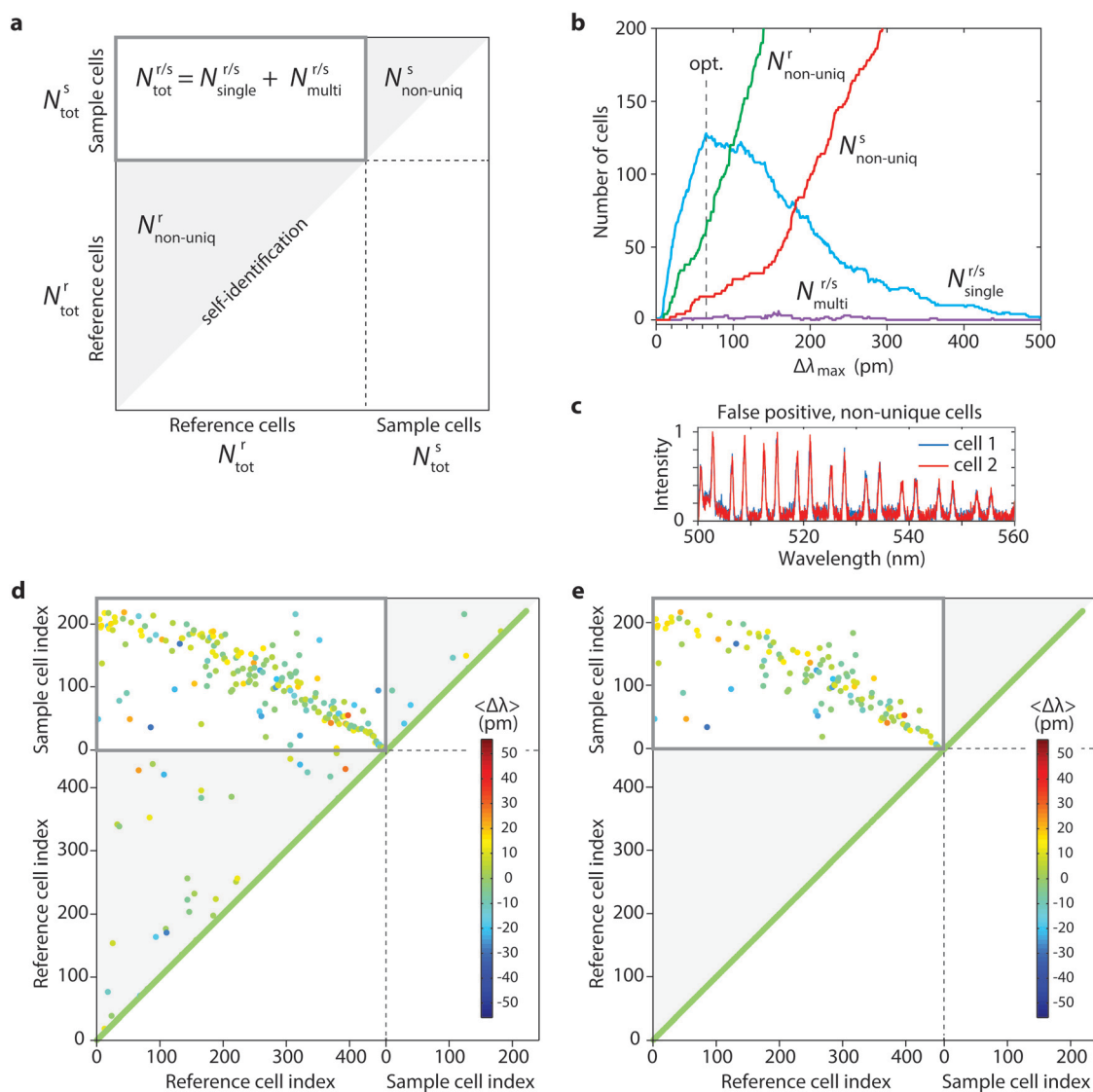


Fig. 4 (a) Representation of a full correlation map. (b) Dependence of identified unique cells and errors on the choice of $\Delta\lambda_{\max}$. An optimum value of 65 pm for $\Delta\lambda_{\max}$ was determined, which maximized the number of identified unique sample cells. (c) An example of two non-unique cells. (d) A correlation map obtained with $\Delta\lambda_{\max} = 65$ pm. (e) A modified correlation map when the non-unique cells in the reference and sample have been removed.

Test with flowing cells in the microfluidic channel

The emission spectra of stationary cells and beads therein were measured to be very stable over repeated measurements. When cells are in motion, however, the intracellular positions of beads and the relative position of multiple beads in a single cell may change, and these movements affect the emission spectra and the barcode and, therefore, may cause errors in barcode identification. We tested the barcode readout scheme for hundreds of cells under non-stationary conditions by making the cells flow in one direction through the microfluidic channel and then flow in the backward direction. In the backward scan, the cells are not expected to be in the exact reverse order because they are mixed in the tubing before returning to the channel. After loading a dispersion of cells into the syringe, the syringe pump was set to a volume flow rate of 250 nl min^{-1} , so that the cells pass the $50 \text{ }\mu\text{m}$ -long laser-excitation zone in the channel, one cell at a time with a speed of 1.67 mm s^{-1} . As the cells flowed through the channel, their emission spectra were acquired at a frame rate of 28 spectra per s. A total of 479 cells containing at least one bead passed through the excitation zone. Of these, a total of 457 ($=N_{\text{tot}}^{\text{f}}$) cells were detected (met the detection criterion) in the forward direction, 320 of which contained a spectral fingerprint from one bead and 137 from more than one bead. An estimation of additional 97 cells did not contain beads, which were thus undetected. All detected cells represent a reference barcode library. Next, we reversed the flow direction and scanned a partial group of the reference cells in the backward flow at the same flow rate. A total of 222 ($=N_{\text{tot}}^{\text{s}}$) cells were detected (*i.e.* met the detection criterion), which constitutes a test sample group.

With the total of 679 ($N_{\text{tot}}^{\text{f}} + N_{\text{tot}}^{\text{s}}$) measured spectra, we drew a full correlation map showing the correlations of 679×679 scan data for a given $\Delta\lambda_{\text{max}}$ (Fig. 4a; ESI†). Each pair of non-unique cells with identical barcodes appears in the lower-left (or upper-right) shaded triangular regions and is counted into $N_{\text{non-uniq}}^{\text{f}}$ (or $N_{\text{non-uniq}}^{\text{s}}$). The optimum $\Delta\lambda_{\text{max}}$ value was chosen, which maximized the number ($N_{\text{single}}^{\text{f/s}}$) of sample cells matched to single reference barcodes. With the measured spectra, a number of various types of correlation counts were calculated as a function of $\Delta\lambda_{\text{max}}$ ranging from 0 to 500 pm (Fig. 4b). At $\Delta\lambda_{\text{max}} = 0$, there are no correlation hits between two different cells (Fig. 4b). The number of unique and non-unique identified cells increases as $\Delta\lambda_{\text{max}}$ increases and the criterion for identification is relaxed. However, as $\Delta\lambda_{\text{max}}$ increases beyond $\sim 65 \text{ pm}$, the number of non-unique cells ($N_{\text{non-uniq}}^{\text{f}}$, $N_{\text{non-uniq}}^{\text{s}}$) continued to increase, and as the non-unique cells are removed from the reference and sample, the number of identified unique cells ($N_{\text{single}}^{\text{f/s}}$) begins to decrease. In principle, the number of non-unique cells, *i.e.* false positive errors, can be reduced by using a batch of beads with all dissimilar diameters. Indeed, some cells appeared to contain similar beads with almost indistinguishable spectra (Fig. 4c). We have chosen $\Delta\lambda_{\text{max}} = 65 \text{ pm}$ that maximized $N_{\text{single}}^{\text{f/s}}$. A full correlation map (Fig. 4d) obtained with $\Delta\lambda_{\text{max}} = 65 \text{ pm}$ shows 403 unique and $54 (=N_{\text{non-uniq}}^{\text{f}})$ non-

unique barcodes in the reference and 206 unique and $16 (=N_{\text{non-uniq}}^{\text{s}})$ non-unique cells in the sample. Without removing the non-unique cells from the reference and sample, 20 sample cells matched with more than one reference cell ($N_{\text{mult}}^{\text{f/s}} = 20$).

To have only unique barcoded cells, the non-unique cells could be physically removed from the reference group. In our experiment, we simulated this situation of pre-sorting in post processing, where all the non-unique cells were removed from the reference and sample groups *in silico* (ESI†). After the removal, a re-calculated correlation map (Fig. 4e) confirms zero false-positive errors in the reference and sample ($N_{\text{non-uniq}}^{\text{f}} = 0$, $N_{\text{non-uniq}}^{\text{s}} = 0$). The cross-correlation map (boxed domain) displays $128 (=N_{\text{single}}^{\text{f/s}})$ identified unique cells and one cell identified multiple times ($=N_{\text{mult}}^{\text{f/s}}$). The remaining $77 (=N_{\text{tot}}^{\text{s}} - N_{\text{non-uniq}}^{\text{s}} - N_{\text{single}}^{\text{f/s}} - N_{\text{mult}}^{\text{f/s}})$ cells in the sample group did not have uniquely matched references (false-negative errors) (ESI†).

Conclusions

The spectral barcode readout scheme is suitable for tagging a few hundreds of cells with reasonable error rates. Most errors stemmed from the overlapping size distribution of current polystyrene beads. Using similar spectral analysis, the beads can be pre-sorted to create a batch with non-overlapping sizes from 8 to $12 \text{ }\mu\text{m}$ with a size increment of 2 nm. Theoretically, two beads that are different in diameter by 2 nm produce WGM barcode peaks that shifted by $|\Delta\lambda_i| = 100 \text{ pm}$ and, therefore, are resolvable when $\Delta\lambda_{\text{max}} = 65 \text{ pm}$. A batch of such uniformly poly-dispersed beads can provide 2000 unique beads and be used to label 2000 cells when each cell is loaded with a single bead. When each cell is loaded with two beads, theoretically, up to 2 000 000 ($= 2000 \times 2000/2$) cells can be uniquely barcoded.²² But this possibility should wait for the development of a method to control the number of beads per cell precisely.

The spectral barcodes are subject to change when the surrounding of beads is altered. Identical beads in environments with different refractive indices can produce non-identical WGM spectra. From simple spectral measurements, we cannot distinguish whether the spectral difference is because of different diameters or same diameters but in different environments. From theory, it can be calculated that the spectral sensitivity to the external refractive index of polystyrene beads is $\sim 90 \text{ nm RIU}^{-1}$. Thus a change in a refractive index of the cytoplasm of 0.001 would cause a spectral shift of $\Delta\lambda \approx 90 \text{ pm}$. In principle, the sensitivity to environment could be reduced by coating beads with a thin layer of low-index material to shield the evanescent field of optical modes from the environment. In addition, the spectral peaks are sensitive to temperature through thermal expansion of beads and temperature-dependent refractive indices. These two factors tend to have opposite effects for most materials. For polystyrene beads, the temperature sensitivity of WGM peaks is calculated to be $3 \text{ pm } ^\circ\text{C}^{-1}$,²² which may be negligible in most applications. Optical cross-talk can arise between beads when

they are in optical contact within the range of evanescent fields and/or when the emission from one bead is absorbed or refracted by its neighboring beads. Nonetheless, we have shown that such cross-talk in uncoated polystyrene beads is small enough to allow more than a hundred unique barcodes, and in principle, the multiplicity of barcodes can be substantially increased by using multiple dyes and 2–3 distinct beads per cell. Further, beads with core–shell structures can minimize optical coupling.

It has been shown that cells containing polystyrene beads with diameters of 7–12 μm undergo cellular division and movement over 1–2 days,^{22–24} and long-term cell viability in the presence of micro- and sub-micron-scale beads has been well established. Nevertheless, the subtle effects of intracellular objects on cellular phenotypes and gene expression patterns are not fully understood and should be an important subject of research in the future. Biocompatibility is expected to improve as beads get smaller, for example, by using high-index materials, such as high-index glass and semiconductors, as well as plasmonics. Also, coating with biocompatible polymers could further enhance biocompatibility.

The polystyrene beads used in this study have a Q-factor of ~ 1000 , a mode spacing of ~ 5 nm, and a finesse of ~ 10 . With this modest finesse, it was proven difficult to distinguish cells containing three or more beads. Fluorescent beads made of materials, such as BaTiO_3 , that have higher refractive indices than polystyrene could be used with the advantage of smaller sizes at the same finesse²⁵ or enhanced Q-factor and finesse at the same size.

Microfluidic chips are useful platforms to scan the barcodes in cells rapidly both for the construction of reference barcode libraries and for the barcode identification of sample cells. In the present study, the barcode scan in the forward flow simulates the process to build a reference barcode library. Test sample sets were produced by sending the cells back to the channel with no or partial cell mixing, and the barcodes of the test cells were compared to the reference library. It is envisioned that in actual applications, sample cells are harvested from cell cultures, tissues, or live animals.

Besides microfluidic chips, the barcode readout method could also be integrated into the conventional flow cytometry platform with modifications to include pumping and high-resolution spectral measurements. State-of-the-art flow cytometers offer 3 to 12 spectral channels.²⁶ The brightness of fluorescent beads is sufficient to enable high-spectral-resolution measurements with millisecond- and sub-millisecond integration times for high-throughput barcode identification and cell sorting. WGM cellular barcoding may prove to be a useful tool for biomedical investigations at single-cell levels. Besides cell tagging, resonator-based barcodes have potential as luminescent probes for highly-multiplexed molecular assays on chip. Finally, the optical barcode technology has potential for non-biological labeling and authentication for documents, work of arts, industrial materials and products.

Acknowledgements

The authors acknowledge discussions with Sheldon Kwok and Nicola Martino. This research was supported by the U.S. National Institutes of Health (R01-CA192878, P41-EB015903, P01-HL120839, DP1-OD022296), National Science Foundation (CBET-1264356, EEC-1358296, ECCS-1505569, CMMI-1562863), and MGH Research Scholar Award. M.H. was supported in part by the Marie Curie International Outgoing Fellowship N^o 627274 within the 7th European Community Framework Programme.

References

- 1 A. M. Klein, L. Mazutis, I. Akartuna, N. Tallapragada, A. Veres, V. Li, L. Peshkin, D. A. Weitz and M. W. Kirschner, *Cell*, 2015, **161**, 1187–1201.
- 2 E. Z. Macosko, A. Basu, R. Satija, J. Nemesh, K. Shekhar, M. Goldman, I. Tirosh, A. R. Bialas, N. Kamitaki, E. M. Martersteck, J. J. Trombetta, D. A. Weitz, J. R. Sanes, A. K. Shalek, A. Regev and S. A. McCarroll, *Cell*, 2015, **161**, 1202–1214.
- 3 E. Lubeck, A. F. Coskun, T. Zhiyentayev, M. Ahmad and L. Cai, *Nat. Methods*, 2014, **11**, 360–361.
- 4 S. C. Bendall, E. F. Simonds, P. Qiu, E. D. Amir, P. O. Krutzik, R. Finck, R. V. Bruggner, R. Melamed, A. Trejo, O. I. Ornatsky, R. S. Balderas, S. K. Plevritis, K. Sachs, D. Pe, S. D. Tanner and G. P. Nolan, *Science*, 2011, **332**, 687–697.
- 5 O. Ornatsky, D. Bandura, V. Baranov, M. Nitz, M. Winnik and S. Tanner, *J. Immunol. Methods*, 2010, **361**, 1–20.
- 6 B. Bodenmiller, E. R. Zunder, R. Finck, T. J. Chen, E. S. Savig, R. V. Bruggner, E. F. Simonds, S. C. Bendall, K. Sachs, P. O. Krutzik and others, *Nat. Biotechnol.*, 2012, **30**, 858.
- 7 Y. Leng, K. Sun, X. Chen and W. Li, *Chem. Soc. Rev.*, 2015, **44**, 5552–5595.
- 8 C. D. Keating and M. J. Natan, *Adv. Mater.*, 2003, **15**, 451–454.
- 9 C. Lin, R. Jungmann, A. M. Leifer, C. Li, D. Levner, G. M. Church, W. M. Shih and P. Yin, *Nat. Chem.*, 2012, **4**, 832–839.
- 10 E. Fernandez-Rosas, R. Gomez, E. Ibanez, L. Barrios, M. Duch, J. Esteve, C. Nogués and J. A. Plaza, *Small*, 2009, **5**, 2433–2439.
- 11 G. Lawrie, B. J. Battersby and M. Trau, *Adv. Funct. Mater.*, 2003, **13**, 887–896.
- 12 Y. Lu, J. Zhao, R. Zhang, Y. Liu, D. Liu, E. M. Goldys, X. Yang, P. Xi, A. Sunna, J. Lu, Y. Shi, R. C. Leif, Y. Huo, J. Shen, J. A. Piper, J. P. Robinson and D. Jin, *Nat. Photonics*, 2013, **8**, 32–36.
- 13 L. C. Mattheakis, J. M. Dias, Y.-J. Choi, J. Gong, M. P. Bruchez, J. Liu and E. Wang, *Anal. Biochem.*, 2004, **327**, 200–208.
- 14 M. Han, X. Gao, J. Z. Su and S. Nie, *Nat. Biotechnol.*, 2001, **19**, 631–635.
- 15 Y. Wang, B. Yan and L. Chen, *Chem. Rev.*, 2013, **113**, 1391–1428.
- 16 K. White, D. Chengelis, K. Gogick, J. Stehman, N. L. Rosi and S. Petoud, *J. Am. Chem. Soc.*, 2009, **131**, 18069–18071.

- 17 J. Livet, T. A. Weissman, H. Kang, R. W. Draft, J. Lu, R. A. Bennis, J. R. Sanes and J. W. Lichtman, *Nature*, 2007, **450**, 56–62.
- 18 D. Cai, K. B. Cohen, T. Luo, J. W. Lichtman and J. R. Sanes, *Nat. Methods*, 2013, **10**, 540–547.
- 19 F. Xiong, N. D. Obholzer, R. R. Noche and S. G. Megason, *PLoS One*, 2015, **10**, e0127822.
- 20 S. A. Dunbar, *Clin. Chim. Acta*, 2006, **363**, 71–82.
- 21 F. Ramiro-Manzano, R. Fenollosa, E. Xifré-Pérez, M. Garín and F. Meseguer, *Adv. Mater.*, 2011, **23**, 3022–3025.
- 22 M. Humar and S. Hyun Yun, *Nat. Photonics*, 2015, **9**, 572–576.
- 23 M. Schubert, A. Steude, P. Liehm, N. M. Kronenberg, M. Karl, E. C. Campbell, S. J. Powis and M. C. Gather, *Nano Lett.*, 2015, **15**, 5647–5652.
- 24 M. Schubert, K. Volckaert, M. Karl, A. Morton, P. Liehm, G. B. Miles, S. J. Powis and M. C. Gather, *Sci. Rep.*, 2017, **7**, 40877.
- 25 M. Humar and S. H. Yun, *Optica*, 2017, **4**, 222–228.
- 26 J. P. Nolan and D. Condello, *Curr Protoc Cytom*, 2013, **63**, 1–18.

Spectral reading of optical resonance-encoded cells in microfluidics

Matjaž Humar,^{1,2,3,†} Avinash Upadhyay^{1,4,†} and Seok Hyun Yun^{1,5,*}

¹Wellman Center for Photomedicine, Harvard Medical School, Massachusetts General Hospital, 65 Landsdowne St. UP-5, Cambridge, Massachusetts 02139, USA

²Condensed Matter Department, J. Stefan Institute, Jamova 39, SI-1000 Ljubljana, Slovenia

³Faculty of Mathematics and Physics, University of Ljubljana, Jadranska 19, SI-1000, Ljubljana, Slovenia

⁴Research School of Engineering, Australian National University, Canberra, North Road Canberra, Australian Capital Territory 0200, Australia

⁵Harvard-MIT Health Sciences and Technology, Cambridge, 77 Massachusetts Avenue Cambridge, Massachusetts 02139, USA

† These authors contributed equally

* syun@hms.harvard.edu

Interpretation of the correlation matrix

In the processing identifying barcodes and comparing cells, many different detection scenarios can be encountered. Here we analyze some of the cases we have observed and illustrate how they presented themselves in the correlation matrix.

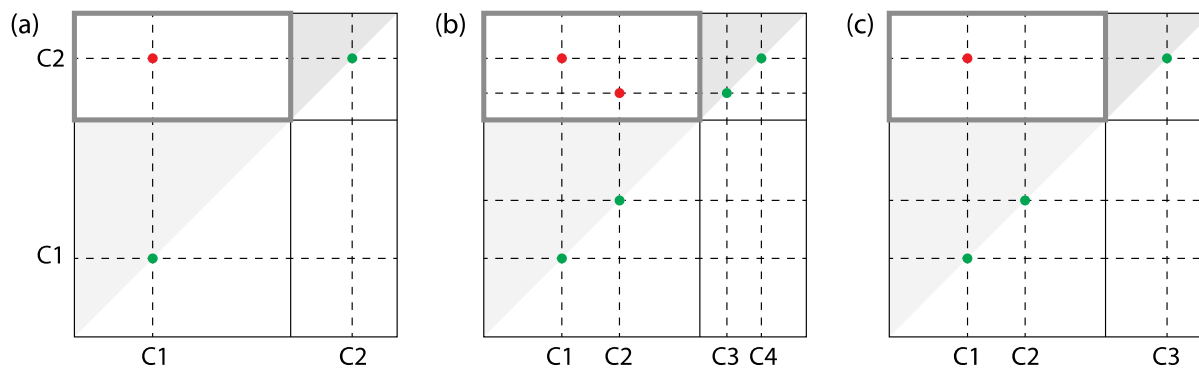


Fig. S1 Normal events. (a) One cell is flown in both directions and successfully identified. (b) Two cells are flown in both directions and successfully identified. (c) Out of two cells flown in the forward direction only one returns.

- (a) In this most common example, for a cell C1 detected in the forward flow and a cell C2 (with its identity yet to be determined) detected in the backward flow, and the protocol decides that these two cells have the same barcode within the maximum spectral shift $\Delta\lambda_{\max}$. Then a correlation mark (red filled circle) is represented in the cross-correlation domain (boxed region), and trivial self-correlation marks (green circles) appear in the reference and sample domain (shaded triangular regions).

- (b) An extended example of the same type as (a): for two cells, C1 and C2, detected in the forward run and two cells, C3 and C4, detected in the backward scan, C3 is identified to be C2, and C4 is identified to be C1, and this make two correlation marks (red) in the cross-correlation domain.
- (c) A cell, C2, detected in the forward scan but not detected in the backward scan. Typically, this situation happens when the cell gets trapped in the tubing or syringe and never returns to the channel.

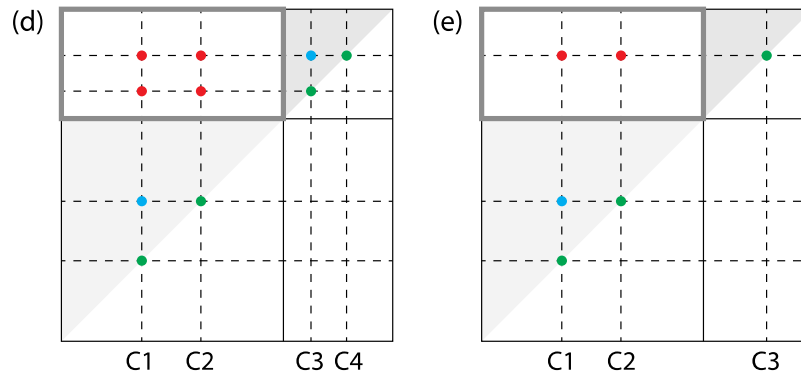


Fig. S2 Scenarios leading to non-unique correlation and false-positive errors: (d) Two cells with identical barcodes are flown in forward direction and both return to the channel in the backward flow. (e) Same as previous example, but only one returns.

- (d) Two cells happen to have very similar beads and their barcodes are considered identical. These cells detected in the forward run, C1 and C2, produces a non-unique correlation hit (cyan-filled circle) in the non-diagonal space (bottom-left area). When these cells return in the backward flow, C3 and C4, they mark a correlation hit in the self-correlation domain (top-right area). In the cross-correlation domain, four correlation marks (red circles) appear.
- (e) Same example as (d) except that now only one of the two cells, C3, return to the channel in the backward scan. Then, C3 matches with both C1 and C2, leaving two correlation marks (red circles).

Note. In the case of (d), unfortunately there is no way to tell whether C3 (or C4) was C1 or C2. And, in (e) we cannot identify whether C3 is C1 or C2. Therefore, in both cases, in post processing we remove all the non-unique cells from reference and sample, by deleting rows and columns C1 through C4 in (d) and rows and columns C1 through C3 in (e), from the dataset. This is how we generated re-processed map in Fig. 4e. In principle it would be enough to remove only cells C1 and C2, which would simulate a sample with only unique-barcodes. However, since barcodes can be altered (Fig. S3) it is better to remove also non-unique cells in the sample to get rid of errors such as in Fig. S3g.

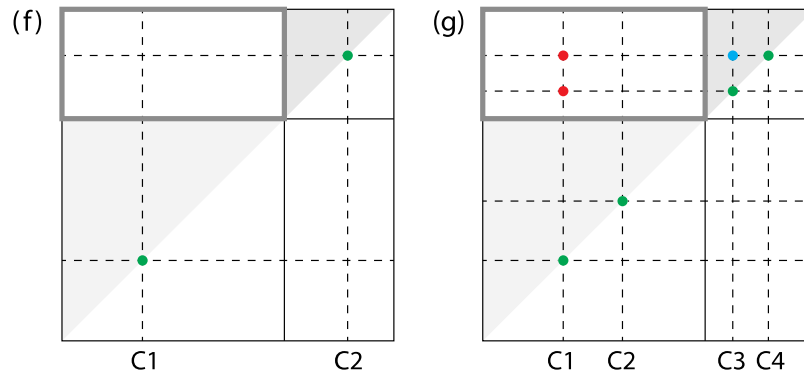


Fig. S3 Scenarios for false-negative errors due to altered barcodes: (f) The barcode of a single cell changes during the reversal of the flow and the cell is not identified. (g) Two unique cells flown in forward direction change their barcode in such way that they become identical.

- (f) The spectrum from the beads in a cell, C2, measured in the backward flow is different from the spectrum of the same cell, C1, measured in the forward flow. This can occur if the effective refractive index of the intracellular environment surrounding the beads has changed between the two measurements. This situation leaves no cross-correlation mark.
- (g) A rare case. The spectra of two cells, C1 and C2, are similar to each other but distinguished in the forward flow, but for any reason, such as cellular or environment changes, the spectra of the two cells, C3 and C4, in the backward direction are changed such that they are no longer distinguished from each other. This makes C3 and C4 non-unique cells. Moreover, suppose C3 and C4 are identified to be the same as C1 but not C2; in this case two correlation marks (red) appear in the cross-correlation domain. C1, C3 and C4 were removed from the dataset when the re-processed map in Fig. 4(e) was generated.