

## Supplementary information for ‘Fluorescence-activated droplet sorting (FADS): efficient cell sorting based on enzymatic activity’

### Model for cellular enrichment by FADS

#### Definition of enrichment

In this manuscript we define the efficiency of a sort in terms of ‘enrichment’: the purification of one cell type (‘positive’ cells) against another (‘negative’ cells). Commonly in FACS enrichment is expressed as follows:

$$\eta' = \frac{N_{+,1}}{N_{-,1} + N_{+,1}} \bigg/ \frac{N_{+,0}}{N_{-,0} + N_{+,0}}$$

where  $\eta'$  is enrichment,  $N_{+,0}$  is the number of positive cells before sorting,  $N_{+,1}$  is the number of positive cells after sorting and  $N_{-,0}$  and  $N_{-,1}$  are the respective values for the negative cells.

An alternative expression of enrichment,  $\eta$ , can be defined as:

$$\eta = \frac{N_{+,1}}{N_{-,1}} \bigg/ \frac{N_{+,0}}{N_{-,0}}$$

This is the expression of enrichment that we have used in the characterization of the sorting device because it enables a quantitative comparison of enrichment over a much wider range of initial and final conditions than  $\eta'$ .

In order to demonstrate the reasoning behind our choice, we present Table S2, which shows the values of  $\eta$  and  $\eta'$  for a set of hypothetical (although typical) sorts. The ‘dynamic range’ of  $\eta'$  is limited and does not allow robust comparisons of enrichment when positive cells outnumber negative cells after sorting; for example, the ten-fold improvement in efficiency seen in case D versus case C results in only a 10% apparent rise in enrichment (11 versus 10). A similar effect is seen in cases G and H (and K and L) where a 10-fold improvement is expressed as a ~2-fold rise in  $\eta'$ . In contrast,  $\eta$  is a good measure of sorting efficiency over a much wider range of enrichments: if the ratio of positive cells to negative cells increases 10-fold after sorting,  $\eta$  rises 10-fold as well, regardless of the initial starting ratio. The only disadvantage to expressing enrichments as  $\eta$  is that when only positive cells are recovered,  $\eta$  rises to infinity. In practice, however, this divergence can be avoided by calculating a finite lower limit for enrichment instead; for example, if an  $\epsilon_0 = 0.1$  cell mixture is sorted and 100 positive cells (and no negative cells) are recovered then  $\eta$  can be recorded as  $>1000$  ( $100/0.1 = 1000$ ), rather than infinity.

#### Model

The droplets are considered as individual compartments in which cells are encapsulated. To begin, only one type of cell is considered. Initially, the cells (total number is  $N_c$ ) are randomly dispersed in a volume  $V$  of nutrient medium. On emulsification, this volume is divided between  $N$  compartments ( $N$  being much larger

than 1). The average number of cells per droplet is  $\lambda = N_c/N$ . Assuming that the cells are randomly distributed, the probability that cell  $i$  is encapsulated in compartment  $j$  is the same for all cells and compartments. Each cell has the probability  $p = 1/N = \lambda/N_c$  to be in compartment  $j$  (or in any other compartment). The probability of having  $k$  cells in a given compartment ( $P(X = k)$ ) corresponds to the probability of encapsulating  $k$  cells (each of them with probability  $p$ ) with the avoidance of  $(N_c - k)$  cells (with a probability  $(1-p)$ ). The probability  $P(X = k)$  is, therefore:

$$\begin{aligned} P(X = k) &= \left(\frac{1}{N}\right)^k \left(1 - \frac{1}{N}\right)^{\lambda N - k} C_{\lambda N}^k \\ &= (N - 1)^{-k} \left(\frac{N - 1}{N}\right)^{\lambda N} \left(\frac{(\lambda N)!}{(\lambda N - k)!k!}\right) \end{aligned}$$

Since  $((N - 1)/N)^N \approx e^{-1}$  and  $\lambda N \gg 1$ ,

$$P(X = k) = \frac{e^{-\lambda}}{k!} \lambda^k$$

The cells are encapsulated according to a Poisson distribution, assuming that the cells are randomly distributed in the nutrient medium<sup>22,28</sup>. Note that if the initial distribution is not random, then the Poisson distribution disappears<sup>29</sup>.

When two populations of cells are considered, positive and negative cells for example, they are assumed to be independent. For each cell type, the probabilities of having  $k_+$  and  $k_-$  are:

$$\begin{aligned} P(X = k_+) &= \frac{e^{-\lambda_+}}{k_+!} \lambda_+^{k_+} \\ P(Y = k_-) &= \frac{e^{-\lambda_-}}{k_-!} \lambda_-^{k_-} \end{aligned}$$

where  $\lambda_+$  and  $\lambda_-$  are the average number of cells (positive or negative) per droplet.

During a complete selection, several rounds of sorting may be performed. For round  $n$ , the ratio of positive cells to negative cells  $\varepsilon_n$  is defined as  $\lambda_+/\lambda_-$ . During the sorting process, positive droplets (which are sorted) will contain at least one positive cell and zero or more negative cells as a result of co-encapsulation (Fig. S4). Assuming that the cells neither proliferate nor die during the experiment, the number of positive and negative cells selected during a sort of  $N$  droplets is:

$$\begin{aligned} N_+ &= N \times \sum_{k_+=1}^{\infty} k_+ \times P(X = k_+) \\ N_- &= N \times \sum_{k_+=1}^{\infty} P(X = k_+) \times \left( \sum_{k_-=1}^{\infty} P(Y = k_-) \times k_- \right) \end{aligned}$$

After the  $n^{\text{th}}$  round of sorting, the ratio of positive to negative cells,  $\varepsilon_n$ , is:

$$\epsilon_n = \frac{N_+}{N_-} = \frac{\lambda_+}{\lambda_-(1 - e^{-\lambda_+})} = \frac{\epsilon_{n-1}}{1 - e^{-\epsilon_n \lambda / (1 + \epsilon_{n-1})}}$$

where  $\lambda = \lambda_+ + \lambda_-$ .

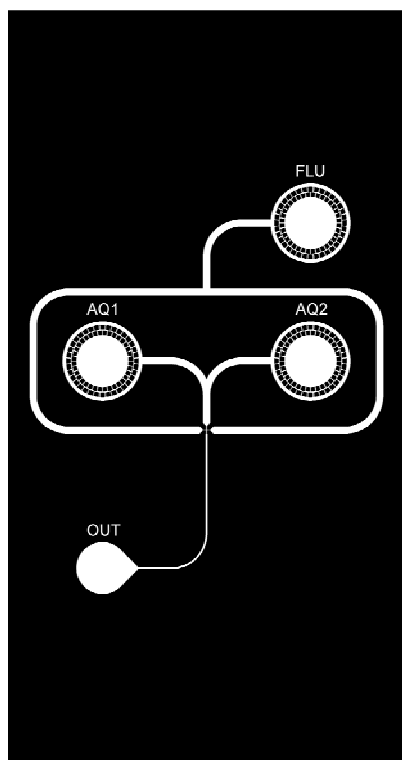
The degree of enrichment achieved for round  $n$ ,  $\eta_n$ , is then obtained as  $\epsilon_n / \epsilon_{n-1}$  (see below).

$$\eta_n = \frac{1}{1 - e^{-\epsilon_{n-1} \lambda / (1 + \epsilon_{n-1})}}$$

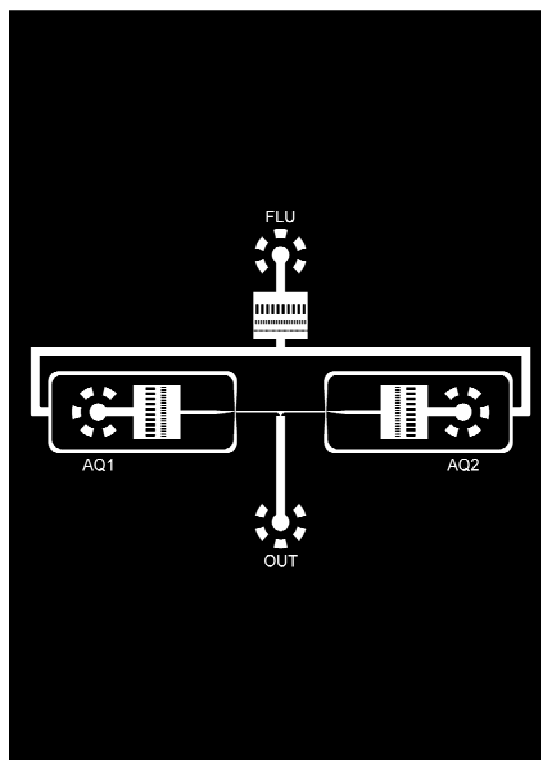
In a typical selection where  $\epsilon_{n-1}$  would be much smaller than 1, a good estimate of the enrichment is obtained using the asymptotic approximation  $\eta \approx 1 / (\epsilon_{n-1} \lambda) \approx 1 / \lambda_+$ . This highlights the fact that enrichment values increase with the dilution of the positive cells.

## Supplementary figures

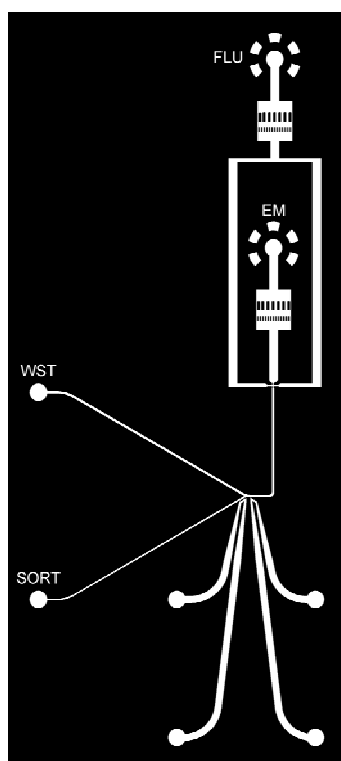
(a)



(b)

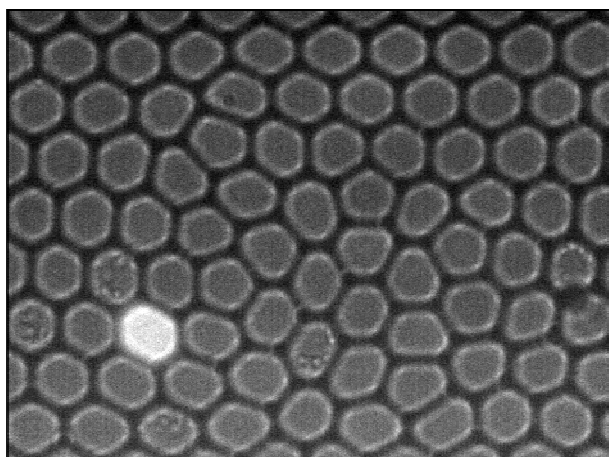


(c)

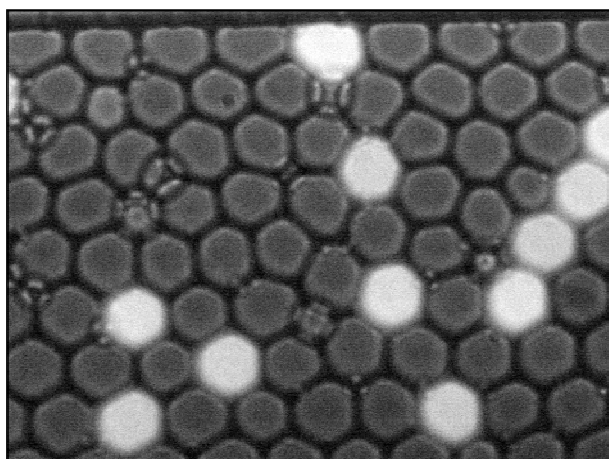


**Fig. S1.** CAD files for the fabrication of the microfluidic devices. (Previous page.) (a) ‘Single-emulsifier’ device. Streams from the two aqueous ports (‘AQ1’ and ‘AQ2’) merge with twin surfactant/fluorinated oil streams from the surfactant/fluorinated oil port (‘FLU’). Droplets are generated by flow-focusing and break-up of the merged aqueous stream at the nozzle ( $20 \times 100 \mu\text{m}$ ). (b) ‘Dual-emulsifier’ device. Two aqueous streams (‘AQ1’ and ‘AQ2’) are individually flow-focused by twin streams of surfactant/fluorinated oil from the surfactant/fluorinated oil port (‘FLU’). (c) Sorting device. 12 pL droplets are injected into the device (‘EM’) and spaced-out with fluorinated oil (‘FLU’). The droplets travel to the junction of the device at a velocity of  $\sim 20 \text{ cm}\cdot\text{s}^{-1}$ . The arm of the junction on the side of the electrodes (the ‘positive arm’ leading to the ‘SORT’ output) is  $40 \mu\text{m}$  wide, while the arm furthest from electrodes (the ‘negative arm’ leading to the ‘WST’ output) is  $60 \mu\text{m}$  wide. The difference in hydraulic resistance between the two arms favors the flow of liquid along the wider channel so that, in the absence of an electric field, all of the droplets follow the main stream along the negative arm. Individual droplets are drawn into the positive arm by electric actuation, triggered on droplet fluorescence. Actuation consists of applying an AC voltage across the electrodes, thereby exerting a dielectrophoretic force on the droplet. Above a voltage threshold (typically  $1\text{--}1.4 \text{ kV}_{\text{p-p}}$ ), this dielectrophoretic force becomes larger than the viscous forces maintaining the droplet in its flow line. The droplet is pulled across the flow lines and flows into the narrow arm of the sorter.

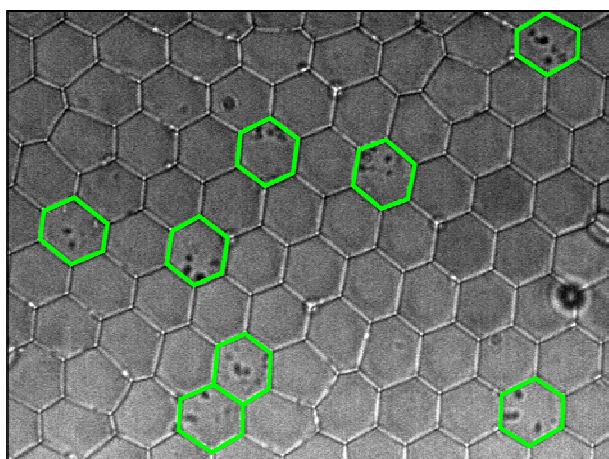
**(a)**



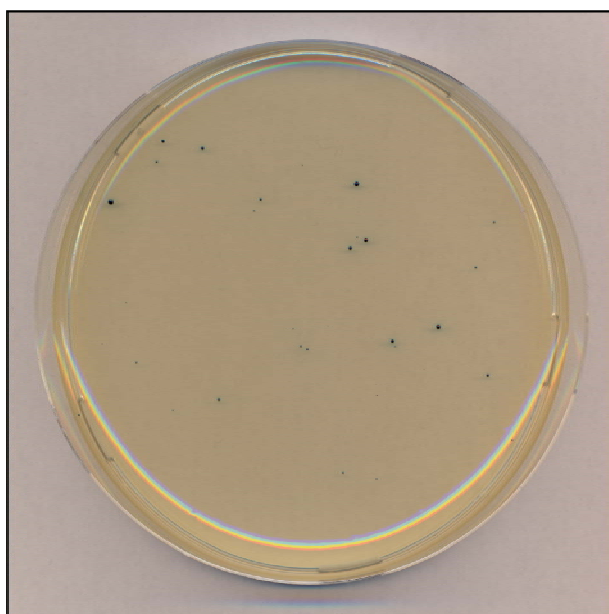
**(b)**



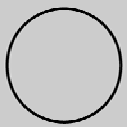
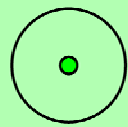
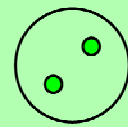
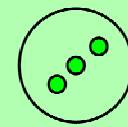
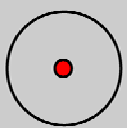
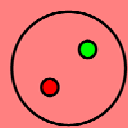
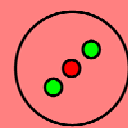
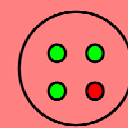

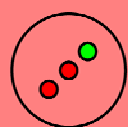
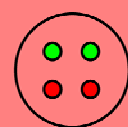
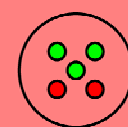
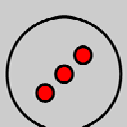
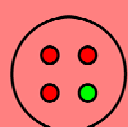
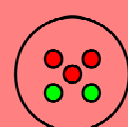
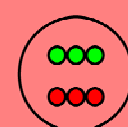
**(c)**



**Fig. S2.** Increasing the fraction of fluorescent droplets by adding a lyzing agent. (Previous page.) (a) Fluorescence micrograph of a  $\lambda = 0.1$  emulsion of *lacZ* cells containing no polymyxin B after overnight incubation. 1.9% of 587 droplets in the emulsion were fluorescent, corresponding to ~19 % of the occupied droplets. The fluorescent droplets contained one or more cells with compromised membranes, allowing contact between cytoplasmic  $\beta$ -galactosidase and the fluorogenic substrate FDG in the droplet. (b) Fluorescence micrograph of a  $\lambda = 0.1$  emulsion of *lacZ* cells containing  $100 \mu\text{g}\cdot\text{ml}^{-1}$  polymyxin B after overnight incubation. Adding polymyxin B caused the cells in all of the occupied droplets to lyse and catalyze the degradation of FDG: 9.9% of 666 droplets were fluorescent, corresponding to ~100% of the occupied droplets. (c) Bright-field micrograph of the emulsion in (a). The occupied droplets were clearly visible (green outlines) owing to the colonies of cells growing inside them. The  $\lambda$  value for this emulsion was confirmed as being ~0.1: 9.4% of 722 droplets contained colonies.



**Fig. S3.** Photograph of *E. coli* colonies recovered from a single sorted droplet. The droplet was sorted from an  $\varepsilon_0 = 1$ ,  $\lambda \approx 0.001$  emulsion which contained  $\sim 3 \times 10^3$  cells ( $3 \times 10^6$  droplets). The sorted droplet was broken into LB medium and spread on LB agar containing ampicillin, IPTG and X-gal. 28 blue colonies and 0 white colonies grew on the plate.

Number of positive cells \ Number of negative cells	Number of positive cells				
	$P(X=0)$	$P(X=1)$	$P(X=2)$	$P(X=3)$	
$P(Y=0)$	 $P(X=0) \times P(Y=0)$	 $P(X=1) \times P(Y=0)$	 $P(X=2) \times P(Y=0)$	 $P(X=3) \times P(Y=0)$	
$P(Y=1)$	 $P(X=0) \times P(Y=1)$	 $P(X=1) \times P(Y=1)$	 $P(X=2) \times P(Y=1)$	 $P(X=3) \times P(Y=1)$	
$P(Y=2)$	 $P(X=0) \times P(Y=2)$	 $P(X=1) \times P(Y=2)$	 $P(X=2) \times P(Y=2)$	 $P(X=3) \times P(Y=2)$	
$P(Y=3)$	 $P(X=0) \times P(Y=3)$	 $P(X=1) \times P(Y=3)$	 $P(X=2) \times P(Y=3)$	 $P(X=3) \times P(Y=3)$	
		Non-fluorescent droplets		Fluorescent droplets	

**Fig. S4.** Limitation of enrichment by co-encapsulation of cells. In each droplet the number of positive cells,  $k_+$ , is given by the statistical distribution  $P(X = k_+)$  and the number of negative cells,  $k_-$ , by  $P(Y = k_-)$ . Since the two distributions are independent, the encapsulation of mixtures of cells is defined by the probabilities of encapsulating each cell type. Droplets containing positive cells are fluorescent (green and red areas). When fluorescent droplets containing only positive cells are sorted, only positive cells are recovered (green area). However, in droplets where negative cells were co-encapsulated with positive cells, sorting leads to co-selection (red area). The overall numbers of positive cells and negative cells in the population of fluorescent droplets determine the enrichment for a given sort.



### Supplementary tables

$\lambda$	$\varepsilon_0$	$\varepsilon_1$ determination (colonies)		
		Blue	White	$\varepsilon_1$
0.021	0.1	213	0	>213
0.016	1	537	0	>537
0.16	0.01	97	72	1.35
0.2	0.1	322	46	7
0.15	1	466	29	16.1
1.6	0.01	46	156	0.294
0.91	0.1	109	197	0.553

**Table S1.** Determination of  $\varepsilon_1$  by colony counting after sorting. This table shows the numbers of blue and white colonies observed growing on LB agar following the recovery of cells from sorted droplets.

Case	No. of cells (initial)		No. of cells (final)		$\eta$	$\eta'$
	Positive	Negative	Positive	Negative		
A	1	10	1	10	1	1
B	1	10	1	1	10	6
C	1	10	10	1	100	10
D	1	10	100	1	1000	11
E	1	100	1	100	1	1
F	1	100	1	10	10	9
G	1	100	1	1	100	51
H	1	100	10	1	1000	92
I	1	1000	1	100	1	10
J	1	1000	1	10	10	91
K	1	1000	1	1	100	501
L	1	1000	10	1	1000	910

**Table S2.** Equivalent values for  $\eta$  and  $\eta'$  in hypothetical sort cases. Note that the numbers of cells merely demonstrate the ratios of cell types before and after sorting: in a real sort the numbers of cells would be much greater.

## Supplementary movie legends

**Movie S1.** High-speed movie of droplet production. 12 pl droplets containing fluorescein were generated in a single-emulsifier device. The aqueous phase was flow-focused by twin streams of fluorinated oil containing a surfactant.

**Movie S2.** High-speed movie of droplet reinjection. 12 pl droplets containing fluorescein were injected into a sorting device. The droplets were spaced-out by twin streams of surfactant-free fluorinated oil.

**Movie S3.** High-speed movie of fluorescence-based sorting of droplets. 12 pl droplets containing either 25 or 100  $\mu\text{M}$  fluorescein were sorted according to fluorescence. The 100  $\mu\text{M}$  droplets were selectively deflected into the positive arm of the sorting junction by dielectrophoresis.

**Movie S4.** High-speed movie of the production of a binary emulsion. 12 pl droplets containing either 25 or 100  $\mu\text{M}$  fluorescein were generated simultaneously using a dual-emulsifier device. Each aqueous phase was flow-focused by twin streams of fluorinated oil containing a surfactant. The droplets were combined and collected at a common exhaust.

**Movie S5.** High-speed movie of cell sorting based on enzymatic activity. 12 pl droplets in an  $\varepsilon = 0.1$ ,  $\lambda = 1.2$  emulsion sorted according to fluorescence. Fluorescent droplets containing *E. coli* pIVEX2.2EM-*lacZ* cells were selectively deflected into the positive arm of the sorting junction by dielectrophoresis.

**Movie S6.** Movie of droplets containing enzymatically-active cells. Three columns of 12 pl droplets in a  $\lambda = 0.52$  emulsion of *E. coli* pIVEX2.2EM-*lacZ* cells passing through a laser line in a microfluidic channel. Approximately 10% of the droplets containing visible colonies were fluorescent. The gray bars at the edge of the image indicate the position of the laser line.