Title: Bacterial Microcolonies in Gel Beads for High-Throughput Screening of Libraries in Synthetic Biology.

Authors: Duarte J.M., Barbier I., Schaerli Y.

Journal: ACS synthetic biology,

Year: 2017

Volume: 6(11)

Pages: 1988-1995

DOI: 10.1021/acssynbio.7b00111

Bacterial microcolonies in gel beads for high-throughput screening of libraries in synthetic biology

José M. Duarte[1]; Içvara Barbier[2] and Yolanda Schaerli[1,2]*

1) Department of Evolutionary Biology and Environmental Studies, University of Zurich, Winterthurerstrasse 190, 8057 Zürich, Switzerland

2) Department of Fundamental Microbiology, University of Lausanne, Biophore Building, 1015 Lausanne, Switzerland

*Correspondence: yolanda.schaerli@unil.ch

Abstract

Synthetic biologists increasingly rely on directed evolution to optimize engineered biological systems. Applying an appropriate screening or selection method for identifying the potentially rare library members with the desired properties is a crucial step for success in these experiments. Special challenges include substantial cell-to-cell variability and the requirement to check multiple states (e.g. being ON or OFF depending on the input). Here, we present a high-throughput screening method that addresses these challenges. First, we encapsulate single bacteria into microfluidic agarose gel beads. After incubation, they harbor monoclonal bacterial microcolonies (e.g. expressing a synthetic construct) and can be sorted according their fluorescence by fluorescence activated cell sorting (FACS). We determine enrichment rates and demonstrate that we can measure the average fluorescent signals of microcolonies containing phenotypically heterogeneous cells, obviating the problem of cell-to-cell variability. Finally, we apply this method to sort a pBAD promoter library at ON and OFF states.

Keywords: Synthetic biology, directed evolution, screening, combinatorial libraries, hydrogel beads, cell-to-cell variability

Introduction

Synthetic networks have potential for numerous biomedical[1] and biotechnological applications[2] and for improving our understanding of nature's design principles[3]. Unfortunately, our ability to rationally design, engineer and predict the behavior of such circuits is still limited[4]. Therefore, it is often necessary to extensively optimize an initial design until a system performs as required. This is commonly done in a time-consuming trial-and-error approach. To make this optimization process more efficient, synthetic biologists increasingly rely on directed evolution[5-11]. Here, genetic diversity is introduced to generate a library containing many variants of a synthetic construct or gene. This library is then subjected to a screen or selection in order to identify the genotype of variants with improved properties. It might be necessary to diversify the selected variants further to create a new library and to repeat the procedure to recover a circuit that works as desired. Possibilities to introduce diversity into a synthetic network are plentiful and facilitated by the advances in DNA synthesis, assembly and cloning methods[12].

In directed evolution, the main challenge is to find the rare functional variants within a library of mutants. For this purpose, a high-throughput screening or selection system is advantageous, since the more library members one can assay, the higher the success rate. In a screening assay, a specific output of the individual library members is measured and appropriate variants are chosen. By contrast, in a selection, the desired behavior of a variant is linked to its survival, so that only the desired mutants will survive the selection. The difficulty for synthetic circuits is that their outputs often have more than one state (e.g. being ON or OFF depending on the input). Consequently, a selection/screening system needs to test all the states in order to identify functional variants. Selection/screening systems that meet these requirements are only just emerging[6] and synthetic biology could still benefit from innovative solutions.

Another complication is that many synthetic networks do not reliably function at the single cell level, resulting in phenotypically heterogeneous cell populations[4]. That means that on average a population of genetically identical cells might display the desired behavior (e.g. a logic gate), but there is a substantial cell-to-cell variability. It is also possible that a proportion of the individual cells do not actually respond to the signal (examples include references *13-15*). This is not of further concern for certain applications, such as a bioreporter[16] where the readout is based on the average signal of several thousand cells. However, this implies that the selection/screening process to obtain this biosensor should ideally also be carried out at the level of cell populations and not at the level of individual cells. This might be a reason why fluorescence activated cell sorting (FACS) – the sorting of single cells according to their fluorescence – has so far not been frequently used for screening libraries of synthetic circuits, despite its high level of flexibility, the possibility to screen multiple parameters at once and the unmatched throughput of $>10^7$ events/hour[17]. It is however possible to make use of the advantages of FACS, while not being dependent on single cell readouts, by sorting small cell colonies instead of single cells. Such microcolonies can be contained in hydrogel beads[18-20] or in water-in-oil-in-water double emulsion droplets[21-23]. Beads with a diameter of approximately up to 50 μm can be sorted by FACS[24]. Bigger beads can be sorted in a large particle sorter such

as the COPAS (Complex Object Parametric Analyzer and Sorter) machine[25-27], although with lower throughput. The microcolonies in the beads are monoclonal, as long as just a single cell per bead/droplet is initially encapsulated, and that cell then grows to a microcolony inside the compartment. Highly monodisperse gel beads are commonly produced by vibrating technology[28] or in water-in-oil emulsion droplets generated on a microfluidic device[29].
Here we demonstrate that hydrogel beads, generated by a two-phase microfluidic device[29] containing *E. coli* microcolonies (expressing a synthetic construct), and their sorting by FACS is a promising and experimentally undemanding high-throughput screening method for synthetic biology. We determine enrichment rates (up to 30,000-fold) and demonstrate that these beads can be used to circumvent the problem of high cell-to-cell variability by measuring the fluorescent signal of the microcolonies. Finally, we use this method to isolate members of a pBAD[30] promoter library with desired fluorescence levels at the ON and OFF states.

Results and Discussion

Method overview
In the first step of our high-throughput screening method (Fig. 1), individual bacteria (*E. coli* in our case) are compartmentalized into monodisperse water-in-oil emulsion droplets[31] using a microfluidic device (Fig. S1). The number of cells per droplet follows a Poisson distribution[32] and can be controlled to ensure single occupancy. The aqueous phase for droplet formation contains the bacteria, agarose and growth medium. By cooling the emulsion on ice, the agarose solidifies inside the droplets and forms gel beads. The entrapped cells grow into microcolonies inside the beads when incubated (typically at 37 °C). Next, the emulsion is broken and the beads are recovered in an aqueous buffer. The beads can then be analyzed and sorted by FACS. Finally, the cells can be recovered from the beads for further analysis, mutagenesis or additional screening rounds.

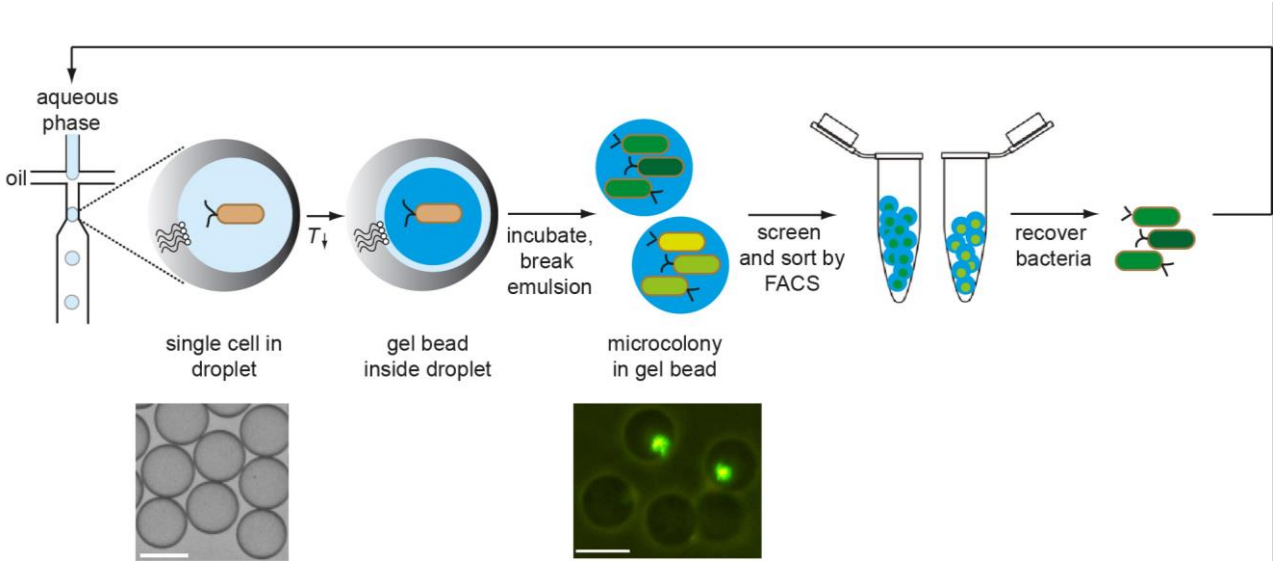

Figure 1. Overview of the screening method. Single cells are encapsulated into monodisperse surfactant-stabilized water-in-oil emulsion droplets produced with a microfluidic device. The

aqueous solution also contains agarose that gelates upon cooling on ice, so that solid gel beads form inside the droplets. During incubation of the emulsion, the cells grow into monoclonal microcolonies inside the beads. The latter are recovered from the emulsion and sorted by FACS. The bacteria are recovered from the gel beads and are then ready for a further round of sorting, mutagenesis or analysis. Pictures: Phase contrast microscope picture of droplets (left) and fluorescence microscope picture of beads with two of them containing a microcolony. Scale bars: 50 µm.

Growth of bacteria in gel beads

First, we confirmed that *E. coli* cells indeed grow to microcolonies inside the beads. To this end, we formed water-in-oil emulsion droplets (diameter 20 µm) with the aqueous phase containing 1% ultra-low melting agarose, the growth medium and *E. coli* cells constitutively expressing superfolder green fluorescent protein (GFP)[33]. After cooling the emulsion on ice, we incubated it at 37 °C and retrieved an aliquot every 2 hours.

Fig. 2a shows scatterplots of the flow cytometry measurements at different time points during incubation. Before incubation, one major population ("A") is observed. It corresponds to beads containing either no cell (80%), a single cell (18%) or two cells (2%; the proportion of single and double cells were estimated using a Poisson distribution[32]). With increasing incubation time, some beads display increased forward scatter (FSC-H) and sideward (SSC-H) values. This appearing second population ("B") corresponds to beads that harbor a microcolony that has grown inside them, which results in increased forward and sideward scatter, and displays high fluorescence (Fig. 2b, gate B). Complementary to this, after 8 h of incubation the population "A" only contains beads without fluorescence (i.e. without cells) and the beads with low fluorescence (i.e. carrying a single cell) disappeared (Fig. 2b, gate A, 8h) as the cells grew to form the microcolonies.

It is important to note, that the variation in fluorescence of the bead population is not increased compared to that of the single cell population. Rather, it decreases (Fig. S2), suggesting that we did not introduce additional variation due to cell growth. If substantial growth differences between microcolonies in different beads existed, we would see them as differences in forward and sideward scatter values, but we do not observe such growth differences in our data (Fig. 2a).

A comparison with the fluorescence of beads loaded with a known amount of fluorescent cells lets us estimate that after 4 h of incubation 20 µm and 50 µm beads contain approximately 30 and 330 *E. coli* cells, respectively (Fig. S3). The achieved cell densities ($>5 \times 10^9$ cells/ml) are comparable to densities previously measured in alginate beads with a diameter of 400 µm[26].

As seen from the fluorescent picture (Fig. 1), the microcolonies occupy only a small volume of the beads. A way to increase cell growth is to break the emulsion before incubation (Fig. S4). The beads can then be incubated in ample growth medium and oxygen leading to higher cell occupancy than when incubated in the emulsion. The disadvantage of this strategy is that all beads share the same medium. It is, for example, not possible to screen for cells that excrete a quorum sensing molecule that then activates a network in the cells from the same bead. The

chemical would also enter neighboring beads, where the molecule is not produced. In this form of incubation, it is also more likely that some cells escape the beads and then grow outside the bead in the growth medium. These free bacteria can then decrease sorting throughput and a filtering step where the beads are separated from the free bacteria might become necessary.

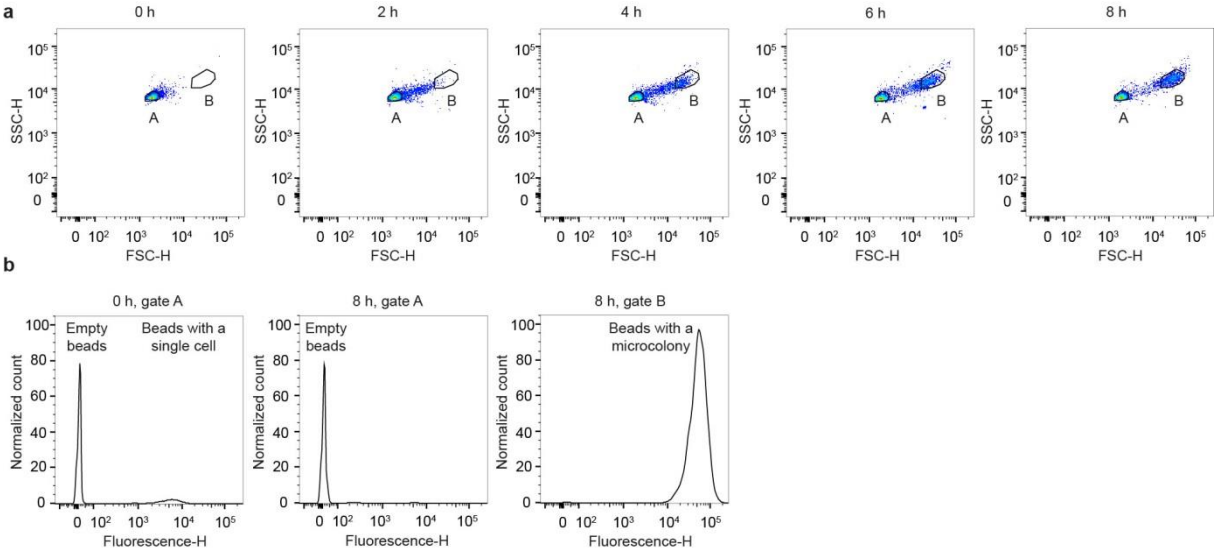

Figure 2. Growth of bacteria to microcolonies inside gel beads. Individual *E. coli* cells constitutively expressing GFP were compartmentalized into gel beads and the emulsion was incubated at 37 °C. Aliquots of the emulsion were broken every 2 hours and the beads were analyzed by flow cytometry. a) Sideward scatter (SSC-H) and forward scatter (FSC-H) of the beads measured by flow cytometry. Gate "A" encircles beads containing no cell or a single cell. Gate "B" encircles the emerging population of beads harboring a bacterial microcolony. b) Fluorescence histograms of beads in gate "A" at 0 h and 8 h and in gate "B" at 8 h.

Determination of enrichment rates

Next, we determined the ability to isolate specific beads from an overwhelming majority of beads with undesired properties. For this enrichment analysis, we mixed *E. coli* cells expressing GFP with cells that do not express GFP at different ratios. We compartmentalized cells (0.3 / droplet) into 20 µm gel beads, and incubated them (4 h) to allow cell growth and GFP expression. Afterwards we sorted them by FACS and collected the fluorescent beads. We liberated the cells by enzymatically digesting the agarose with the enzyme agarase and plated them out on agar plates. Subsequently, we regrew a representative sample of the recovered colonies and analyzed them by flow cytometry. We determined the enrichments by dividing the percentage of GFP positives after sorting by that before sorting (Table 1). For a 1:100,000 dilution (0.001%) of GFP-expressing cells we achieved an enrichment of 30,000-fold. The enrichment rates are comparable to enrichment rates previously reported for the sorting of double emulsions by FACS[21-23]. Such high enrichment rates are possible because of the high throughput of FACS.

Table 1. Enrichment of fluorescent cells in gel beads.

% of GFP-positive cells before sorting	% of GFP-positive cells after sorting	enrichment (-fold)
1	76	76
0.1	90	900
0.01	22	2,200
0.001	30	30,000

Ability to find averages of bimodal signals using population measurements

Next, we wanted to confirm that measuring the fluorescence of microcolonies in beads can be advantageous as compared to measuring the fluorescence of single cells. For this purpose, we transformed the *E. coli* strain TOP10 with a plasmid where the expression of GFP is under the control of the arabinose-inducible pBAD promoter[30]. In TOP10 cells, genes controlled by the pBAD promoter are expressed in such a way that the percentage of expressing cells in the population correlates with the arabinose concentration, and not the degree of expression in individual cells[34]. The underlying all-or-none gene expression is due to the arabinose-dependent expression of the arabinose transporter[34] and results in a phenotypically heterogeneous population with a bimodal distribution (Fig. 3a). We encapsulated TOP10 cells harboring the pBAD-GFP plasmid into gel beads (50 µm diameter), incubated them at different arabinose concentrations and measured their fluorescence by flow cytometry (Fig. 3b). The observed fluorescence levels of the microcolonies are no longer bimodally distributed, indicating that we measure the average fluorescence of a phenotypically heterogeneous population and the beads can therefore be used to perform screenings on the population level.

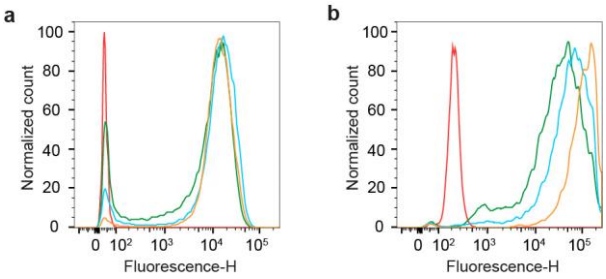

Figure 3. Ability of beads to yield the average signal of a phenotypically heterogeneous population. *E. coli* cells (TOP10) harbor a plasmid where the GFP expression is controlled by the arabinose-inducible pBAD promoter. Red: 0% (w/v) arabinose, Green: 0.0004% (w/v) arabinose, Blue: 0.0005% (w/v) arabinose, orange: 0.001% (w/v) arabinose. a) Cells were grown in solution and individual cells were measured by flow cytometry. The expression is an all-or-none event in individual cells, resulting in a bimodal distribution. b) Cells were encapsulated in gel beads, grown to microcolonies and the beads were analyzed by flow cytometry.

Sorting a pBAD library

Encouraged by these results, we went on to use this method to sort a library of pBAD promoter mutants cloned upstream of the GFP reporter (pBAD-GFP) in TOP10 cells.

In TOP10 cells, mutations in the pBAD promoter can change the ratio of the ON and OFF subpopulations, as well as the positions of the ON and OFF peaks (Fig. S5). Therefore, only measurements based on the population-average can capture all the different behaviors of the library members.

To generate a library, we completely randomized the sequence between the -10 and -35 boxes and the three flanking nucleotides up- and downstream, respectively[35], which ensures that the most important functional elements (-10 and -35 boxes) are left unchanged. Fig. 4 shows the fluorescence distribution of the library compared to the starting wild-type (WT) construct as free cells (Fig 4a) as well as microcolonies in gel beads (50 µm diameter) (Fig 4b).

The vast majority of mutants have a decreased fluorescence when compared to the WT construct. The average fluorescence of the library population is 50% of the WT (Fig. 4c). We collected beads with lower fluorescence than WT (ON, “low”, 700 beads) and higher fluorescence (ON, “high”, 200 beads) at medium induction level (0.0005% arabinose). Subsequently, we performed a second sort to remove mutants with high 'leaky' expression without induction. To this end, we re-encapsulated the sorted cells into beads with 0% arabinose (OFF) and isolated the microcolonies without fluorescence. Fig. 4c shows the corresponding changes of the fluorescence of the populations in the expected directions. This demonstrates that it is possible to FACS-sort heterogeneous populations in gel beads based on their average behavior.

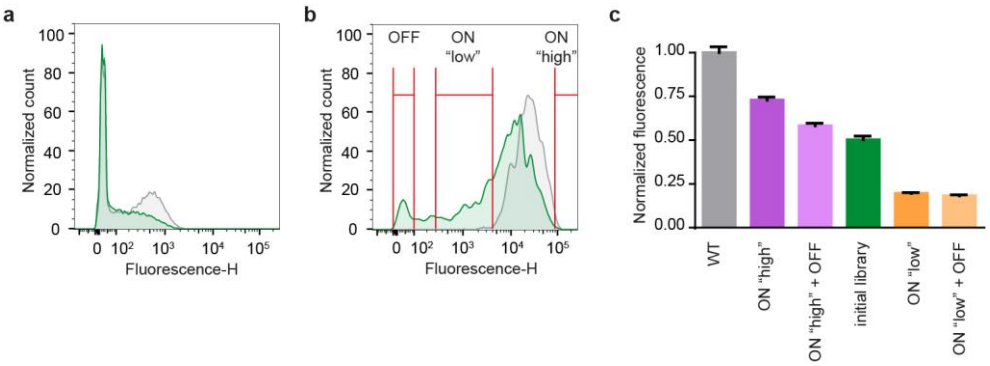

Figure 4. Sorting of a library. Green: randomized pBAD promoter library, grey: WT pBAD promoter. a) Cells were grown in solution with 0.0005% (w/v) arabinose and individual cells were measured by flow cytometry. b) Cells were encapsulated in gel beads with 0.0005% (w/v) arabinose, grown to microcolonies and the beads were analyzed by flow cytometry. The red regions indicate the gates used during the sorts. Note that the OFF sort was performed with 0% arabinose and the ON sorts at 0.0005% (w/v) arabinose. c) Average GFP fluorescence of the libraries before and after the indicated sorts normalized to the fluorescence of pBAD(WT)-GFP at 0.01% arabinose. Error bars represent standard deviations of three replicates.

Finally, we characterized four pBAD mutants from both ("high" and "low") sorts spanning a wide range of activities, by measuring their fluorescence in the absence (0%) and presence (0.01%) of arabinose and sequencing them (Figs. 5, S5 and S6). The highest level of fluorescence measured at 0% arabinose was only 1.5-fold higher than WT, suggesting that the sort to avoid leaky promoters was effective. The highest fluorescence level measured at high induction (0.01% arabinose) was 125% of the WT fluorescence. This is in good agreement with previous low-throughput screenings of pBAD libraries randomizing this promoter region[36]. If the goal is to identify mutants with even higher activity, our library-generation strategy should be modified. For example, we could introduce less mutations (e.g. by error-prone PCR[37]), while at the same time targeting the whole promoter sequence, including the -10 and -35 boxes.

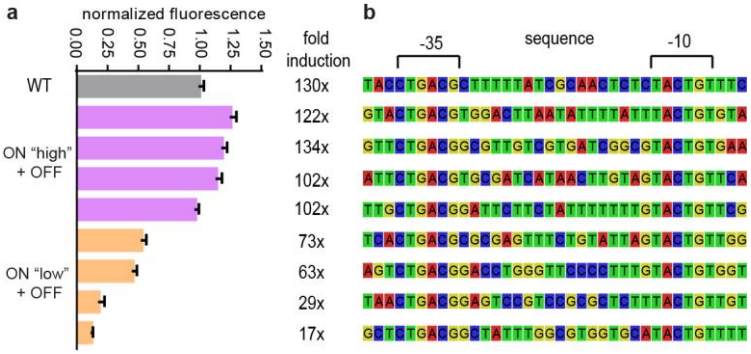

Figure 5. Characterization of selected mutants from the sort. a) GFP fluorescence of pBAD promoter mutants normalized to the fluorescence of pBAD(WT)-GFP at 0.01% arabinose. Error bars represent standard deviations of three biological replicates. Fold inductions are the ratios between the GFP fluorescence at 0.01% and 0% arabinose. b) Sequences of the pBAD promoter mutants. The -35 and -10 boxes were kept constant.

In summary, we demonstrate here that growing monoclonal microcolonies in gel beads and sorting them by FACS is a promising high-throughput method for screening libraries in synthetic biology. This approach is especially useful if the screening must be based on an average signal of a phenotypically heterogeneous population. Sorting a population instead of single cells is also advantageous if the signal to be detected is very low, because the signal can be amplified by growing the cells in a bead. This is for example useful for the directed evolution of enzymes, if combined with a strategy to co-compartmentalize the enzyme reaction product in the bead[24].

A complementary approach to obtain an average population signal for each mutant in a bimodal population could be to pool data from multiple single-cell measurements. However, such an approach is impractical when a large library of different variants must be analysed. Each mutant must be identified, separated, and measured individually in order to later pool the data. In contrast, our approach allows us to conveniently quantify the average population signal of all library members simultaneously in the same pool. However, for other types of cell-to-cell

variability than bimodality repeatedly sampling each library member multiple times as single cell might provide an alternative approach to the one presented here.

The work-flow outlined here produces highly monodisperse beads and the sorting by FACS is technically less demanding than sorting droplets on a microfluidic device[38] or producing double emulsion droplets[21, 22]. FACS also has a higher throughput and is accessible to more researchers than large particle sorters such as the COPAS[25-27]. It is possible to sort for variants that take on different states (e.g. ON and OFF) under different conditions (e.g. by varying inducer concentrations or incubation times) by performing multiple sequential rounds of sorting. This method could also be easily amended to screen for other readouts than the expression of a fluorescent protein. It is possible to assay cell growth by staining microcolonies with a fluorescent biomass indicator dye (e.g. staining nucleic acids or proteins)[18]. Another option is to co-encapsulate sensor cells that fluoresce upon the production of a compound of interest[25]. When encapsulating multiple cells with different genotypes per droplet, cell-cell interactions, for example for distributed computing[39], could also be screened for. Thus, the described method is broadly applicable in synthetic biology.

Methods

Gel-bead formation

Poly(dimethyl)siloxane (PDMS) microfluidic devices were purchased from Wunderlichips GmbH. The devices (see Supplementary Figure S1) contain two inlets: one for the aqueous phase and the other one for the fluorinated oil HFE-7500 (3M Novec) with 0.5% (w/w) 008-FluoroSurfactant (RAN Biotechnologies). Two versions of the device were used: i) 16 µm channel width at the flow focusing part and 20 µm channel height. With this device flow rates were typically 60 µl/h for the aqueous phase and 1400 µl/h for the oil phase, resulting in droplets with a diameter of about 20 µm produced at a frequency of around 4000 kHz. ii) 40 µm channel width at the flow focusing part and 50 µm channel height. With this device flow rates were typically 80 µl/h for the aqueous phase and 2000 µl/h for the oil phase, resulting in droplets with a diameter of about 50 µm produced at a frequency of around 340 kHz. The solutions were hold in glass syringes (100 µl and 5 ml Hamilton Gastight syringes) and connected to the device with PTFE tubing (1/32"ID x 1/16"OD, Cole Parmer) and steel couplers (20 ga x 15 mm; Instech Laboratories). The syringe with the aqueous phase was kept warm by covering it with a warm hot/cold compress. The device was run on an inverted microscope (10x objective) connected to a DALSA GENIE-HM640 fast camera. The droplets were harvested off-chip in Eppendorf tubes.

Cell encapsulation

E. coli cultures (from overnight incubations or from defrosted glycerol stocks) were spun down and resuspended in fresh medium and the absorbance was measured (NanoDrop). To correlate the measured absorbance to the number of bacteria in a sample, initially a standard curve of serial dilutions was generated and plated for viable counts. The measured absorbance was

plotted versus the colony forming units (CFU) for each *E. coli* strain used. Encapsulation of *E. coli* into droplets follows the Poisson distribution[32]. To obtain mainly monoclonal compartments, suspensions were prepared that contained maximal 0.3 bacteria per droplet, resulting in 74%, 22% and 3% of the beads containing none, one, or two cells, respectively. Cells were diluted with medium and mixed 1:1 with 2% agarose (Type IX-A, Ultra-low Gelling Temperature, Sigma) dissolved in the same medium. Arabinose was added when indicated. Beads were prepared as described above.

Cell growth

The emulsion was cooled on ice for at least 15 min. It is also possible to incubate the emulsion directly after generation – without prior bead formation by cooling. We did not observe any differences in growth of the cells between the two options. We chose to keep the samples on ice until all were prepared and then incubate them all at the same time. The emulsion was overlaid with light mineral oil (Sigma-Aldrich) to prevent evaporation and incubated at 37 °C for the indicated amount of time. After incubation, the emulsion was again cooled on ice for 15 min. The fluorinated oil was removed as much as possible. 500 µl phosphate buffered saline (PBS) were added and then 20-50 µl 1H,1H,2H,2H-perfluorooctanol (PFO, Sigma-Aldrich) were added to break the emulsion. The tube was shortly vortexed (5 s) and briefly centrifuged (10 s, 2500 g). The aqueous phase containing the beads was then transferred to a clean tube. Before flow cytometry analysis or sorting all samples were filtered (50 µm mesh, CellTrics, Partec).

Flow cytometry data analysis

FlowJo (FlowJo, LLC) was used for flow cytometry data analysis. We found that plotting the width of the sideward scatter (SSC-W) versus the height of the forward scatter (FSC-H) is the combination of scatter parameters that separates the beads best from other particles (free bacteria, agarose debris, dust). We therefore gated first all beads (with and without microcolonies) on the SSC-W – FSC-H dot-plot. For separating beads with and without microcolonies we used the SSC-H – FSC-H dot plot. The data is displayed on bi-exponential scales.

Plasmids

The plasmid scaffolds and the cloning procedure have been described previously[40]. pET-J23100-TetO-GFP-LVA constitutively expresses GFP. It is identical to KM229387 in the GenBank nucleotide database, except that it has a Tet operator (TetO) instead of a Lac operator. However, the function of TetO was not used in this study. pCOLA-AraC-pBAD-GFP expresses GFP under the control of the pBAD promoter and pCOLA-AraC-TetR-LVA-UmuD-SP6RNAP (KM 229380) was used as non-fluorescent control plasmid.

Time-course experiment (Fig. 2)

E. coli cells (TOP10, Invitrogen) were transformed with a plasmid that expresses GFP constitutively (pET-J23100-TetO-GFP-LVA). Luria-Bertani medium (LB) was supplemented

with 100 μg/ml ampicillin. 20 μm droplets were generated as described above, aliquoted into several tubes and incubated at 37 °C for the indicated amount of time. The beads were recovered as described above and analysed on a BD LSRFortessa (488 nm laser, 530/30 nm filter) flow cytometer.

Enrichment experiments (Table 1)

E. coli cells (MK01[41]) were transformed with pCOLA-AraC-pBAD-GFP (resulting in fluorescent cells) or pCOLA-AraC-TetR-LVA-UmuD-SP6RNAP (resulting in non-fluorescent cells). The LB medium supplemented with 50 μg/ml kanamycin and 0.1% L-arabinose (Sigma-Aldrich) was used. The fluorescent and non-fluorescent *E. coli* cells were mixed to give the ratios as indicated in Table 1. 20 μm droplets were formed as described above. The emulsion samples were incubated at 37 °C for 4 h. Beads were recovered as described above and sorted using a BD FACSAria III (488 nm laser, 530/30 nm filter, 70 μm nozzle): Beads containing *E. coli* microcolonies were gated in the log FSC-H vs SSC-W plot. Beads from this population were then further gated according their fluorescence in the log fluorescence-H vs log FSC-H plot and beads with high fluorescence (>400 positive beads per sample) were collected into a tube with 250 μl PBS.

To recover the bacteria from the 20 μm beads, the beads were diluted in PBS to a density of approximately 5000 beads/ml PBS. 5 units of β-agarase I (NEB) were added per 1 ml PBS and incubated at 42 °C for 30 min. The bacteria were plated out on a LB agar plate and incubated overnight at 37 °C. On average, we recovered 65% of colonies from the number of beads collected. The colonies were used to inoculate 5 ml LB (supplemented with 50 μg/ml kanamycin and 0.1% L-arabinose) and grown for 5.5 h at 37 °C (225 rpm shaking). The *E. coli* cells were then analysed on a BD FACSCanto (488 nm laser, 530/30 nm filter) flow cytometer. After gating the *E. coli* cells the ratio between the fluorescent and non-fluorescent cells was determined. The enrichment[21] was calculated as $E = \frac{\%\ positives\ after\ sorting}{\%\ positives\ before\ sorting}$.

Averaging the signal (Fig. 3)

E. coli cells (TOP10, Invitrogen) carrying the plasmid pCOLA-AraC-pBAD-GFP were encapsulated in 50 μm beads as described above. M63 medium supplemented with 0.2% glycerol, 0.025% casamino acids, 0.00005% thiamine, 50 μg/ml kanamycin and 0%, 0.0004%, 0.0005% or 0.001% (w/v) arabinose as indicated was used. The emulsion was incubated overnight at 37 °C, the beads were recovered as described above and analysed on a BD LSRFortessa (488 nm laser, 530/30 nm filter). The controls with free cells were treated the same way, just leaving out the step of encapsulation into beads.

pBAD library generation

The megaprimer method[42] was used to generate the pBAD library. Oligonucleotides were ordered from Microsynth (Table 2). Polymerase Chain Reactions (PCRs) were carried out with KOD Hot Start polymerase (MERCK MILLIPORE). PCR products were purified using the QIAquick PCR Purification Kit (QIAGEN). A 165 bp megaprimer library was produced by PCR (95 °C 2 min, 25 cycles of 95 °C 30 s, 45 °C 1 min, 70 °C 10 s) using the forward randomized

oligonucleotide primer pBAD_lib_s and reverse primer Seq_r and pCOLA-AraC-pBAD-GFP as template. In a second PCR, the megaprimer library (475 ng) was used to amplify the whole plasmid pCOLA-AraC-pBAD-GFP (95 °C 2 min, 25 cycles of 95 °C 30 s, 45 °C 1 min, 70 °C 2 min). The template plasmid was digested with DpnI (New England BioLabs) and the PCR was purified. 25 ng were transformed into electrocompetent TOP10 *E. coli* cells (Invitrogen). After recovery (1 ml SOC medium, 1 h, 37 °C) 50 µl were plated out on LB-agar plate (supplemented with 50 µg/ml kanamycin) and incubated overnight (37 °C) to determine the library size (~10^5). The remaining cells were added to 20 ml LB (supplemented with 50 µg/ ml kanamycin), grown overnight (shaking, 37 °C) and glycerol stocks (15%) were prepared and stored at -80 °C. Sanger sequencing (Microsynth) of selected mutants was performed with the primer pBAD_f.

Table 2. Oligonucleotide sequences

Name	Sequence (5' -> 3')
pBAD_lib_s	GATTAGCGGTTCCNNNCTGACGNNNNNNNNNNNNNNNNN TACTGTNNNTCCATACCGAATTC
Seq_r	GAAAGCTGGTCCAAGCGATTG
pBAD_f	GCCGTCACTGCGTCTTTTAC

Sorting of the pBAD library

E. coli cells (TOP10, Invitrogen) carrying the pBAD library were encapsulated in 50 µm beads as described above ("*Averaging the signal"*) and sorted on a BD FACSAria III (488 nm laser, 530/30 nm filter, 100 µm nozzle). The first sort was for being ON with 0.0005% (w/v) arabinose either for high or low fluorescence (see gates in Fig. 4b). At least 200 beads per condition were collected into LB. The beads were directly plated out on pre-warmed LB-agar plates (supplemented with 50 µg/ ml kanamycin) and incubated overnight at 37 °C.

For the 50 µm beads we recovered >10x more colonies than the number of beads collected. The colonies from the plates were resuspended in LB (50 µg/ml kanamycin) and stored as glycerol stocks (15%) at -80 °C. The glycerol stocks were thawed and the cells were again encapsulated into beads to perform a second sort for being OFF at 0% arabinose.

Measurements of the average fluorescence

To measure the fluorescence of the libraries (Fig. 4c) glycerol stocks were defrosted. To measure the fluorescence of the individual mutants (Fig. 5a) overnight cultures of three biological replicates were grown in LB. The cultures were washed and diluted to OD_{600} 0.1 in M63 supplemented with 0.2% glycerol, 0.025% casamino acids, 0.00005% thiamine, 50 µg/ml kanamycin. 120 µl were added to the wells of a 96-well plate together with arabinose (2.4 µl) at 3 different concentrations (0%, 0.0005% or 0.01%). The absorbance at 600 nm and the green fluorescence (excitation: 485 nm, emission: 535 nm) were measured every 10 min in a Tecan Infinite F200 plate reader (37 °C, 220 s orbital (2mm) shaking between readings) until the *E. coli* cells reached stationary phase (data shown at 10 h). Both absorbance and fluorescence were

background corrected. The fluorescence was then normalized for the number of cells by dividing by the absorbance.

Supporting Information

Design of the microfluidic device used for droplet generation; variation in fluorescence of free cells and of microcolonies in beads; estimation of number of cells per bead; incubation of beads in emulsion versus incubation in medium; flow cytometry analysis of individual pBAD promoter mutants; pBAD promoter mutants in a strain with homogeneous expression of the pBAD promoter

Author Information

Corresponding Author
*E-mail: yolanda.schaerli@unil.ch

Address: Department of Fundamental Microbiology, University of Lausanne, Biophore building, 1015 Lausanne, Switzerland.

ORCID
Yolanda Schaerli: 0000-0002-9083-7343

Author Contributions

J.M.D. performed experiments and analyzed data. I.B. performed experiments and analyzed data. Y.S. designed and performed experiments, analyzed data and wrote the paper.

Notes

The authors declare no competing financial interest.

Acknowledgments

We thank Dirk van Sway from Wunderlichips GmbH for device production, Ljiljana Mihajlovic for performing plate reader experiments and Mark Isalan, Javier Santos Moreno and Kathleen Sprouffske for critical reading. The authors acknowledge the assistance and support of the Flow Cytometry Facility, University of Zurich for performing the bead sorting experiments. Funding: Swiss National Science Foundation, Ambizione award (PZ00P3_148235) and University Research Priority Program "Evolution in Action: From Genomes to Ecosystems" at the University of Zurich

References

[1] Xie, M., Haellman, V., and Fussenegger, M. (2016) Synthetic biology-application-oriented cell engineering, *Curr Opin Biotechnol 40*, 139-148.

[2] Smanski, M. J., Zhou, H., Claesen, J., Shen, B., Fischbach, M. A., and Voigt, C. A. (2016) Synthetic biology to access and expand nature's chemical diversity, *Nat Rev Microbiol 14*, 135-149.

[3] Lim, W. A., Lee, C. M., and Tang, C. (2013) Design principles of regulatory networks: searching for the molecular algorithms of the cell, *Mol Cell 49*, 202-212.

[4] Kwok, R. (2010) Five hard truths for synthetic biology, *Nature 463*, 288-290.

[5] Tizei, P. A., Csibra, E., Torres, L., and Pinheiro, V. B. (2016) Selection platforms for directed evolution in synthetic biology, *Biochem Soc Trans 44*, 1165-1175.
[6] Schaerli, Y., and Isalan, M. (2013) Building synthetic gene circuits from combinatorial libraries: screening and selection strategies, *Mol BioSyst 9*, 1559-1567.
[7] Cobb, R. E., Sun, N., and Zhao, H. (2013) Directed evolution as a powerful synthetic biology tool, *Methods 60*, 81-90.
[8] Porcar, M. (2010) Beyond directed evolution: Darwinian selection as a tool for synthetic biology, *Syst Synth Biol 4*, 1-6.
[9] Dougherty, M. J., and Arnold, F. H. (2009) Directed evolution: new parts and optimized function, *Curr Opin Biotechnol 20*, 486-491.
[10] Bassalo, M. C., Liu, R., and Gill, R. T. (2016) Directed evolution and synthetic biology applications to microbial systems, *Curr Opin Biotechnol 39*, 126-133.
[11] Brodel, A. K., Jaramillo, A., and Isalan, M. (2016) Engineering orthogonal dual transcription factors for multi-input synthetic promoters, *Nat Commun 7*, 13858.
[12] Gillam, E. M. J., Copp, J. N., and Ackerley, D. (2014) *Methods in Molecular Biology: Directed Evolution Library Creation*, Vol. 1179, Humana Press, Totowa, NJ.
[13] Elowitz, M. B., and Leibler, S. (2000) A synthetic oscillatory network of transcriptional regulators, *Nature 403*, 335-338.
[14] Auslander, S., Auslander, D., Muller, M., Wieland, M., and Fussenegger, M. (2012) Programmable single-cell mammalian biocomputers, *Nature 487*, 123-127.
[15] Lou, C., Liu, X., Ni, M., Huang, Y., Huang, Q., Huang, L., Jiang, L., Lu, D., Wang, M., Liu, C., Chen, D., Chen, C., Chen, X., Yang, L., Ma, H., Chen, J., and Ouyang, Q. (2010) Synthesizing a novel genetic sequential logic circuit: a push-on push-off switch, *Mol Syst Biol 6*, 350.
[16] van der Meer, J. R., and Belkin, S. (2010) Where microbiology meets microengineering: design and applications of reporter bacteria, *Nat Rev Microbiol 8*, 511-522.
[17] Davies, D. (2012) Cell separations by flow cytometry, *Methods Mol Biol 878*, 185-199.
[18] Weaver, J. C., Bliss, J. G., Powell, K. T., Harrison, G. I., and Williams, G. B. (1991) Rapid clonal growth measurements at the single-cell level: gel microdroplets and flow cytometry, *Bio/technology 9*, 873-877.
[19] Sahar, E., Nir, R., and Lamed, R. (1994) Flow cytometric analysis of entire microbial colonies, *Cytometry 15*, 213-221.
[20] Zengler, K., Toledo, G., Rappe, M., Elkins, J., Mathur, E. J., Short, J. M., and Keller, M. (2002) Cultivating the uncultured, *Proc Natl Acad Sci U S A 99*, 15681-15686.
[21] Zinchenko, A., Devenish, S. R., Kintses, B., Colin, P. Y., Fischlechner, M., and Hollfelder, F. (2014) One in a million: flow cytometric sorting of single cell-lysate assays in monodisperse picolitre double emulsion droplets for directed evolution, *Anal Chem 86*, 2526-2533.
[22] Terekhov, S. S., Smirnov, I. V., Stepanova, A. V., Bobik, T. V., Mokrushina, Y. A., Ponomarenko, N. A., Belogurov, A. A., Jr., Rubtsova, M. P., Kartseva, O. V., Gomzikova, M. O., Moskovtsev, A. A., Bukatin, A. S., Dubina, M. V., Kostryukova, E. S., Babenko, V. V., Vakhitova, M. T., Manolov, A. I., Malakhova, M. V., Kornienko, M. A., Tyakht, A. V., Vanyushkina, A. A., Ilina, E. N., Masson, P., Gabibov, A. G., and Altman, S. (2017) Microfluidic droplet platform for ultrahigh-throughput single-cell screening of biodiversity, *Proc Natl Acad Sci U S A*.
[23] Aharoni, A., Amitai, G., Bernath, K., Magdassi, S., and Tawfik, D. S. (2005) High-throughput screening of enzyme libraries: thiolactonases evolved by fluorescence-activated sorting of single cells in emulsion compartments, *Chem Biol 12*, 1281-1289.
[24] Fischlechner, M., Schaerli, Y., Mohamed, M. F., Patil, S., Abell, C., and Hollfelder, F. (2014) Evolution of enzyme catalysts caged in biomimetic gel-shell beads, *Nat Chem 6*, 791-796.

[25] Meyer, A., Pellaux, R., Potot, S., Becker, K., Hohmann, H. P., Panke, S., and Held, M. (2015) Optimization of a whole-cell biocatalyst by employing genetically encoded product sensors inside nanolitre reactors, *Nat Chem 7*, 673-678.
[26] Walser, M., Leibundgut, R. M., Pellaux, R., Panke, S., and Held, M. (2008) Isolation of monoclonal microcarriers colonized by fluorescent E. coli, *Cytometry. Part A : the journal of the International Society for Analytical Cytology 73*, 788-798.
[27] Walser, M., Pellaux, R., Meyer, A., Bechtold, M., Vanderschuren, H., Reinhardt, R., Magyar, J., Panke, S., and Held, M. (2009) Novel method for high-throughput colony PCR screening in nanoliter-reactors, *Nucleic Acids Res 37*, e57.
[28] Whelehan, M., and Marison, I. W. (2011) Microencapsulation using vibrating technology, *Journal of microencapsulation 28*, 669-688.
[29] Theberge, A. B., Courtois, F., Schaerli, Y., Fischlechner, M., Abell, C., Hollfelder, F., and Huck, W. T. (2010) Microdroplets in microfluidics: an evolving platform for discoveries in chemistry and biology, *Angew Chem Int Ed Engl 49*, 5846-5868.
[30] Guzman, L. M., Belin, D., Carson, M. J., and Beckwith, J. (1995) Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter, *J Bacteriol 177*, 4121-4130.
[31] Tawfik, D. S., and Griffiths, A. D. (1998) Man-made cell-like compartments for molecular evolution, *Nat Biotechnol 16*, 652-656.
[32] Huebner, A., Srisa-Art, M., Holt, D., Abell, C., Hollfelder, F., deMello, A. J., and Edel, J. B. (2007) Quantitative detection of protein expression in single cells using droplet microfluidics, *Chem Commun (Camb)*, 1218-1220.
[33] Pedelacq, J. D., Cabantous, S., Tran, T., Terwilliger, T. C., and Waldo, G. S. (2006) Engineering and characterization of a superfolder green fluorescent protein, *Nat Biotechnol 24*, 79-88.
[34] Siegele, D. A., and Hu, J. C. (1997) Gene expression from plasmids containing the araBAD promoter at subsaturating inducer concentrations represents mixed populations, *Proc Natl Acad Sci U S A 94*, 8168-8172.
[35] Hammer, K., Mijakovic, I., and Jensen, P. R. (2006) Synthetic promoter libraries--tuning of gene expression, *Trends Biotechnol 24*, 53-55.
[36] DTU Denmark iGEM 2011. The araBAD promoter (BBa_I0500).
[37] Alper, H., Fischer, C., Nevoigt, E., and Stephanopoulos, G. (2005) Tuning genetic control through promoter engineering, *Proc Natl Acad Sci U S A 102*, 12678-12683.
[38] Baret, J. C., Miller, O. J., Taly, V., Ryckelynck, M., El-Harrak, A., Frenz, L., Rick, C., Samuels, M. L., Hutchison, J. B., Agresti, J. J., Link, D. R., Weitz, D. A., and Griffiths, A. D. (2009) Fluorescence-activated droplet sorting (FADS): efficient microfluidic cell sorting based on enzymatic activity, *Lab Chip 9*, 1850-1858.
[39] Regot, S., Macia, J., Conde, N., Furukawa, K., Kjellen, J., Peeters, T., Hohmann, S., de Nadal, E., Posas, F., and Sole, R. (2011) Distributed biological computation with multicellular engineered networks, *Nature 469*, 207-211.
[40] Schaerli, Y., Munteanu, A., Gili, M., Cotterell, J., Sharpe, J., and Isalan, M. (2014) A unified design space of synthetic stripe-forming networks, *Nat Commun 5*, 4905.
[41] Kogenaru, M., and Tans, S. J. (2014) An improved Escherichia coli strain to host gene regulatory networks involving both the AraC and LacI inducible transcription factors, *J Biol Eng 8*, 2.
[42] Sarkar, G., and Sommer, S. S. (1990) The "megaprimer" method of site-directed mutagenesis, *Biotechniques 8*, 404-407.