SUPPLEMENTARY INFORMATION

Detailed description of microfluidic devices

Device #1 is a simple coflow droplet maker which created droplets containing yeast or *E. coli* cells in their growth medium (**Supplementary Fig. 5** and **Supplementary Movie 1**). The channels in this device were 25μm tall. The droplets formed by this device were 75μm in diameter. There were three inputs into Device #1: cells, media, and a fluorinated oil/surfactant mixture. The cells injected into this microfluidic device originated from a shake flask culture at an OD600 ~ 0.2 so that all of the cells were in the exponential growth phase. The cells were then centrifugated (at 500×g for yeast and 3,250xg for *E. coli*) for 15 minutes, resuspended in PBS, and this process was repeated three times to remove all remnants of the supernatant. The reinjected cell solution contained 0.025% (w/v) xantham gum to prevent cells from settling in the input syringe. The media was at a 2× concentration since it would be diluted in the device. The 1× concentration of xylose was 5g/L for the yeast experiments and 60g/L and for the *E. coli* lactate fermentation. The distribution of cells in a droplet follows a Poisson distribution where the probability of finding a droplet with *k* cells follows the equation:

$$\mathbf{Prob}(k;\lambda) = \frac{e^{-\lambda}\lambda^{k}}{k!} \qquad \text{Equation 1}$$

where λ is the average number of cells per droplet. Supplementary Figure 1 shows the distribution of cells in droplets based on two cell concentrations. At a yeast cell density of OD = 0.05 (λ = 0.19) in the syringe containing the cells, 16% have 1 cell, 83% are empty, and 1% have 2 cells. For an OD = 0.15 (λ = 0.54), 30% of the droplets have 1 cell, 60%

are empty, and 10% have 2 cells. Thus, the cell density can be varied to ensure that either the number of droplets with one cell is maximized which allows for higher throughput or that there is a lower percentage of droplets with more than one cell which reduces the amount of error in the system. In our yeast experiments, we chose an incoming cell density which was a compromise between the two. We chose an incoming cell density of OD = 0.003 for our *E. coli* experiments. The droplets formed in the droplet maker were collected in a syringe and cultured microaerobically by capping the syringe (**Supplementary Fig. 6 and 7**).

The cells were cultured for a predetermined amount of time in an incubator at 30°C for the yeast cells or 37°C for the *E. coli* cells. Then, the droplets from the syringe were reinjected into Device #2 which contained modules for assay droplet formation, droplet coalescence, fluorescence detection, and sorting (**Supplementary Fig. 8**). The channels in this device were 75μm tall. Two of the inputs into this device were for droplet reinjection and additional oil to space the droplets properly. The other two inputs were for assay droplet production. The reinjected and assay droplets flowed in an alternating sequence through the channel leading up to the coalescence module (**Supplementary Movie 2**). The assay droplets had a diameter of 225μm which was larger than the 75μm reinjected droplets. Due to the parabolic velocity profile in the channels, the reinjected droplets flowed faster than the assay droplets so that the droplet pairs were in contact before reaching the coalescence electrodes. These electrodes applied an AC potential of 1kV at a frequency of 20kHz to destabilize the droplet interface which caused the droplets to combine (**Supplementary Movie 3**)[1]. When using this device, it

was more preferable to have an excess of assay droplets because uncoalesced assay droplets only slightly reduced the throughput while having excess reinjected droplets produced measurement errors because an assay droplet would coalesce with two reinjected droplets.

After coalescence, the droplets flowed through long microfluidic channel delay lines to allow the assay reaction to proceed. After traveling through the channels for 30 seconds, a blue laser spot at a wavelength of 488nm was placed in the middle of the channel to excite the fluorescent dye, resorufin. This dye emitted light at an orange wavelength and was detected by a photomultiplier tube with a filter centered at 593nm. A software program written with LabView was used to analyze the detection data so that the maximum value of the resorufin fluorescence intensities could be recorded[2].

The sorting portion of the device contained two outlet channels. The upper channel, which contained the "desired" droplets, had a constriction to produce a higher hydrodynamic resistance. As a result, the droplets naturally flowed into the lower "undesired" channel. Droplets only flowed into the "desired" droplet channel when the detection system measured a fluorescence value within a predetermined range. At that time, a 2kV AC pulse at 900Hz was applied through the electrodes. The resulting AC field created an electrical potential gradient across the channel which dielectrophoretically moved the droplet towards the electrodes and into the "desired" droplet channel (**Supplementary Movie 4**)[3]. The contents of the two channels were collected using 1mL unfiltered pipet tips. The resulting emulsion was separated into two distinct phases using a fluorinated alcohol, and the aqueous phase, which contained the

cells, was added to a liquid or solid culture to grow the cells. If the droplets flowing through the channel leading up to the sorting portion of the device had insufficient spacing, the “undesired” droplets flowed into the “desired” channel. Optimizing the flow rates of the reinjected and assay droplets controlled the spacing between droplets, but for a large spacing, the pairs of droplets needed a longer time to come in contact with each other in the coalescence module. As a result, a long channel was placed before the coalescence electrodes in this design (**Supplementary Fig. 8**).

Calculation of false positive error rate

There are two types of general errors from the sorting system: false negatives and positives. The number of false negatives is essentially zero because the system is designed to ensure that all sorted droplets are pulled electrophoretically by the sorting electrodes. False positives are caused by droplets which should flow into the undesired channel but flow into the desired channel because the pressure differential between the channels is not sufficiently high or because droplets are too close together as they approach the sorting area of the device. **Supplementary Equation 2** was utilized to quantify the false positive sorting error using data from the xylose enrichment experiments where the initial ratio of H131/TAL1 was 1:1 (**Supplementary Table 1**). The desired and undesired droplets contain cells from the high and low xylose consuming strains, respectively.

$$\frac{Dz + x\left(y_D - Dz\right)}{Uz + x\left(y_U - Uz\right)} = A \qquad \text{Equation 2}$$

Where:

- x is the false positive error rate as defined as the % of droplets which randomly go into the "desired" channel
- y is the % of incoming desired (D) or undesired (U) droplets (out of the total number of droplets)
- z is the % of droplets intentionally sorted
- D is the % of intentionally sorted desired droplets (out of the total number of intentionally sorted droplets)
- U is the % of intentionally sorted undesired droplets (out of the total number of intentionally sorted droplets)
- A is the measured enrichment of desired vs. undesired droplets

In these calculations, the intentional sorting rate and random error rate are assumed to be independent. The values for D and U are from the experiments where the H131 and TAL1 strains were grown separately. Furthermore, growth contributions from both the screening system (e.g., cell growth in droplet) and from the cell preparation (e.g., 5mL tube and 50mL shake flask culturing) were also included since A is based on the number of droplets instead of the number of cells.

The numerator and denominator in **Supplementary Equation 2** are the actual number of desired and undesired droplets, respectively. The Dz term is the percentage of desired droplets out of the entire droplet population. The $x(y_D - Dz)$ term is the percentage of desired droplets which normally would not be sorted but are because of the

false positive error rate. The terms in the denominator are similar except that they are for the undesired droplets.

Xylose assay reaction in droplets

To test the Amplex UltraRed xylose assay reaction in droplets flowing through delay lines, the microfluidic device shown in **Supplementary Figure 9** was used. The channels in this device were 75μm tall. This device mixes two aqueous inputs, one containing xylose and another containing the assay mixture. Droplets were formed when the aqueous stream, a mixture of xylose and the assay, came in contact with the oil stream. Then, the droplets flowed through long microfluidic channel delay lines to allow the assay reaction to proceed. This device was designed so that fluorescence could be measured at different locations along the delay line to determine the optimal point.

Several experiments were run using this device in which the xylose concentration was varied. The fluorescence is well correlated with the xylose concentration (**Supplementary Fig. 10**).

Advantages of microfluidic droplet screening method over traditional techniques

Traditionally, high xylose consumption is screened by serially subculturing an incoming cell population, typically through 10 or more rounds which often takes a week or more to complete. Our high-throughput screening platform by comparison needs only a single round of culturing plus 2 to 3 hours of screening time to enrich a population, and the time needed for additional rounds of screening only increases additively. Furthermore, serial

subculturing is limited to discovering strains which grow faster. Strains which consume xylose faster do not necessarily grow more quickly, and our system could detect such mutants.

We demonstrated in our genomic library screening experiment that our method is far more efficient than the traditional serial subculturing technique for identifying high xylose consuming strains. Two weeks of serially subculturing in xylose liquid medium was necessary to identify strains which had higher growth and consumed more xylose. In comparison, less than a week was necessary using the microfluidic droplet screening method. Furthermore, we also observed that after retransforming the plasmids isolated from the serial subculturing process into a clean background, the rate of recovering the high xylose consumption phenotype was two times lower compared with our method. We hypothesize that this is due to the production of background mutations during the much longer serial subculturing time required in the absence of the microfluidic system.

The results from screening a xylose transporter library showed how our method can identify strains which consume more xylose but do not grow faster. This library consisted of an H131 background with a pRS425GPD plasmid. The plasmid contained the gene shuffling library of two genes: *Debaryomyces hansenii* XM_458532 and *Candida glabrata* strain CBS138 XM_444845. These two genes were selected by using BLAST to search for genes with the most DNA sequence homology to *S. cerevisiae* gene HXT7, the native transporter with the highest xylose uptake affinity[4, 5]. The resulting yeast library was grown on glucose-containing agar plates and the appropriate amino acid dropouts for selection and 1000 colonies formed.

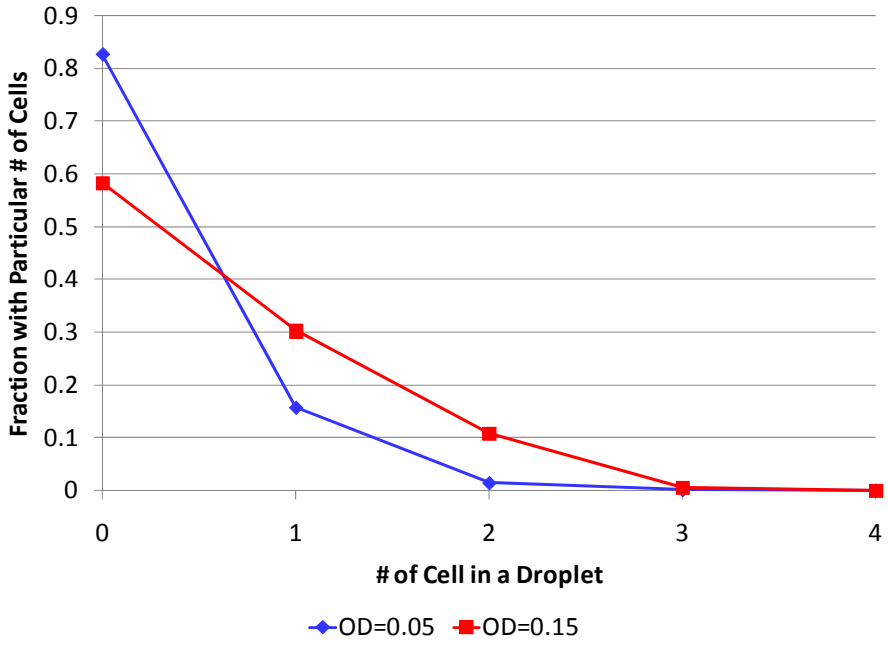
Two rounds of selection were performed, and mutant 2-3 was identified as highest xylose consuming strain. This clone consumed more xylose than the library as a whole but its growth was identical to that of the library. **Supplementary Figure 11** shows that the growth of the mutant is identical to that of the library while **Supplementary Figures 12 and 13** show a consistent difference between the mutant and the library as a whole during mid-exponential phase (e.g., between 24 to 40 hours). At the 40-hour time point, mutant 2-3 had a 4.9% higher xylose consumption than the library, and this difference was statistically significant. These graphs were produced using data from biological replicate experiments. These results show that the microfluidic system has the sensitivity to select for clones with a slightly higher xylose consumption even when their growth is identical to the entire library. In contrast, using the serial subculturing and plating methods did not succeed in identifying mutants with higher xylose uptake from this library since they are limited to finding clones with improved growth phenotypes.

Assay cost advantages over 96 well plate technology

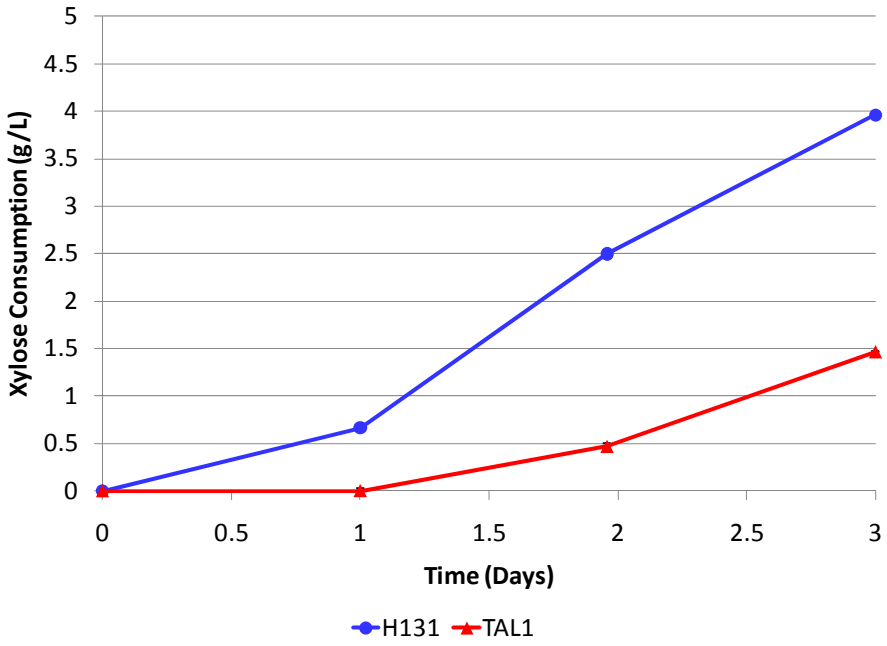
The volume of the droplets formed in these devices is approximately 1nL, which is considerably smaller than the 50μL per well needed with 96 well plates. As a result, the cost of the assay chemicals needed for high-throughput screening is decreased dramatically. For the example reported in this work, the reagent cost to measure the xylose concentration in 10^4 clones was reduced by more than 2000 fold relative to the reagent cost when using 96 well plates. 1mL of 2.6mg/mL Amplex UltraRed, 100U/mL pyranose oxidase, and 10U/mL HRP costs $4.16. Performing one round of screening to

sort 10^4 clones using 96 well plates costs $2078 because 10^4 wells with 50μL of assay solution is 500mL. On the other hand, using droplets costs $4.16. In this case, the volume used is 1mL which is the practical volume used to fill a 1mL syringe even though 2×10^4 droplets with 3nL of assay solution is 60μL.

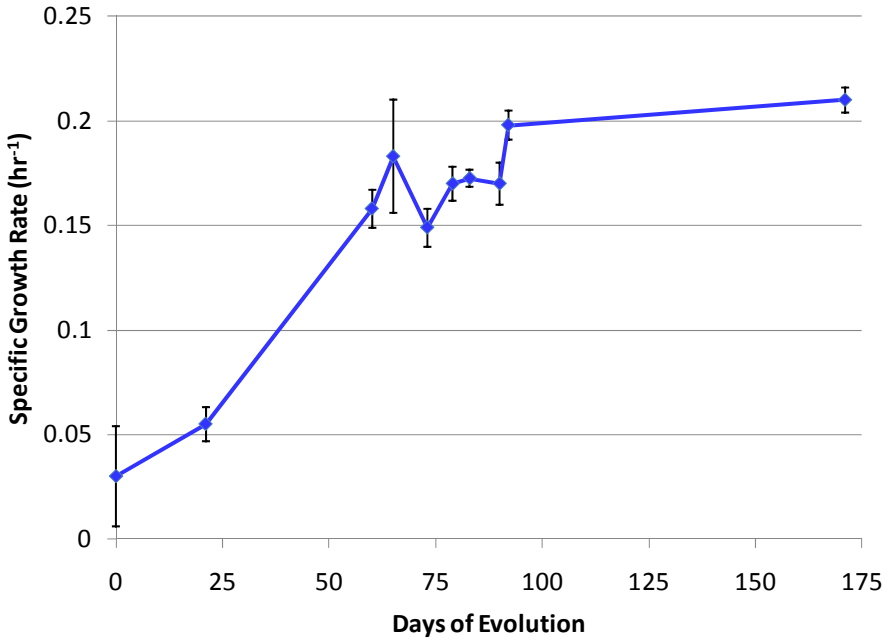
Supporting Figures



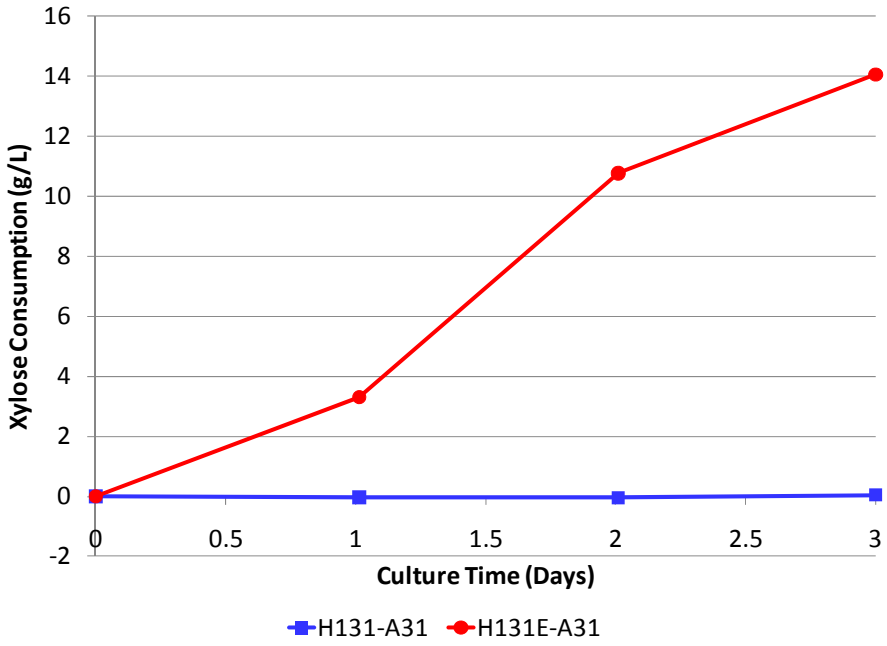
Supplementary Figure 1 Poisson distribution of cells in droplets at two incoming yeast cell densities



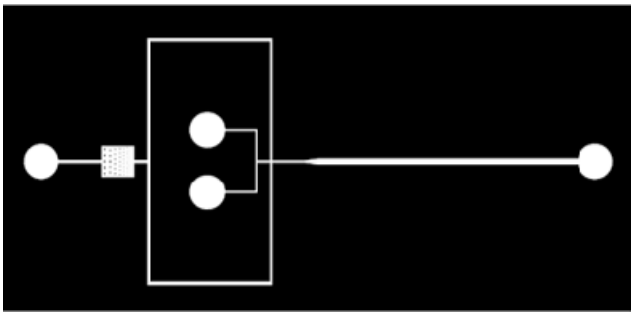
Supplementary Figure 2 HPLC measurements of H131 and TAL1 xylose consumption from 25mL micraerobic fermentation over the period of 3 days which shows that H131 consumes xylose at a higher rate than TAL1



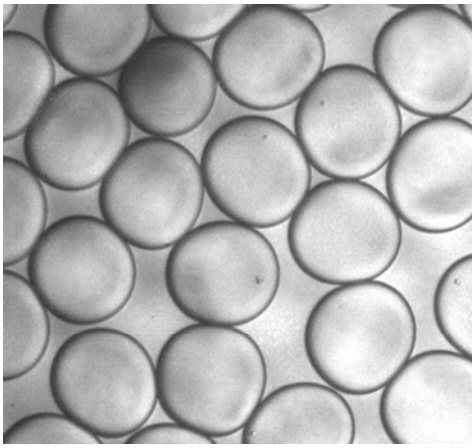
Supplementary Figure 3 Evolution of H131-A31 strain over the period of 6 months. The initial growth rate was below 0.05 hr^{-1} and stabilized around 0.2 hr^{-1} after 90 days. Error bars denote the standard deviation from biological replicate shake flask experiments (n = 2).



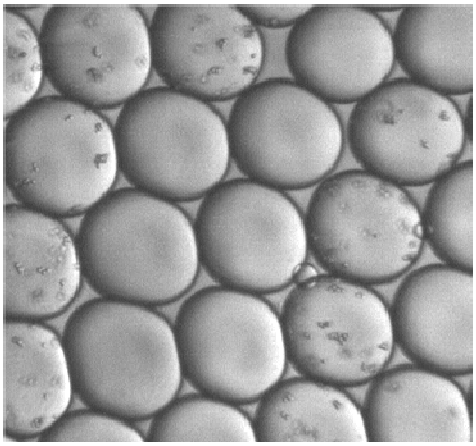
Supplementary Figure 4 Xylose consumption of H131-A31 and H131E-A31 strains. H131-A31 consumed negligible amounts of xylose while H131E-A31 consumed 14g/L xylose after 3 days.



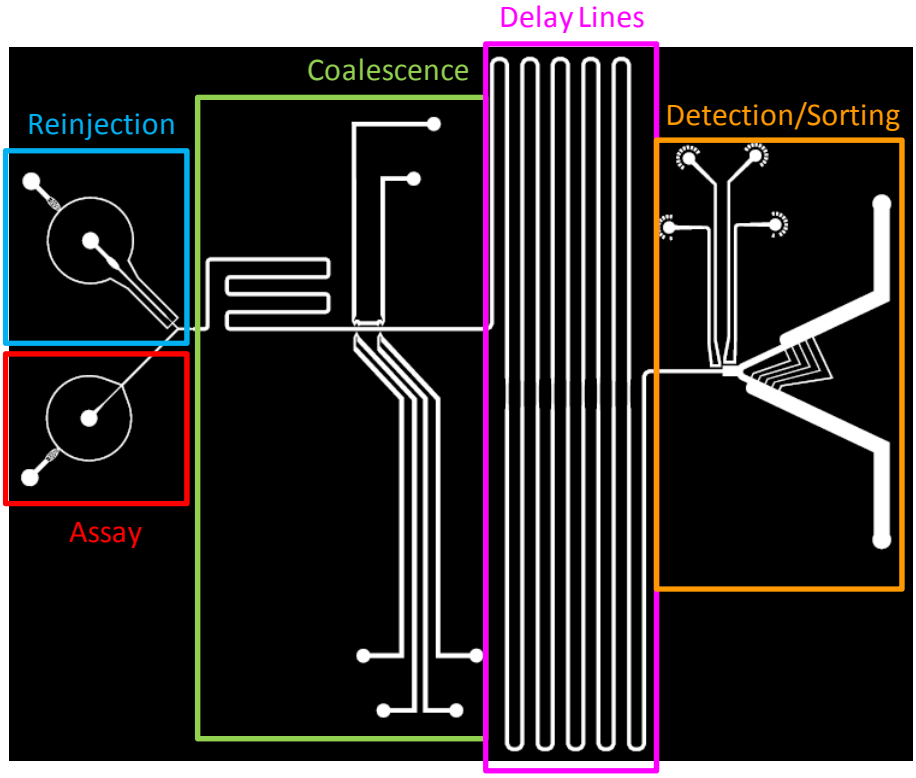
Supplementary Figure 5 Mask design for droplet making device which contains 3 inputs for the oil, cell culture medium, and the cells, and one output.



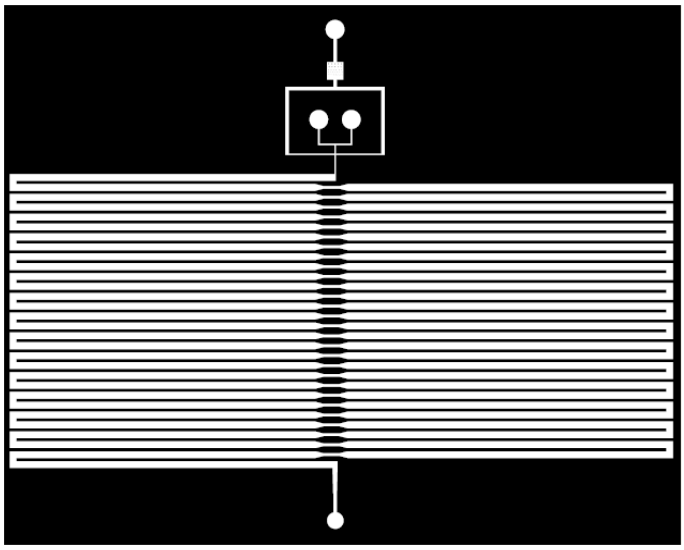
Supplementary Figure 6 Single cells in droplets immediately after collection



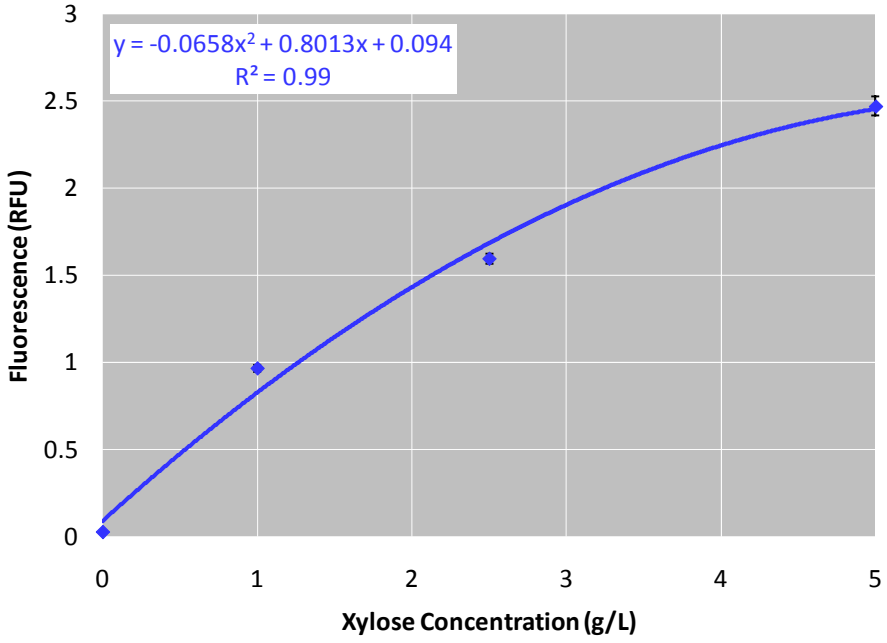
Supplementary Figure 7 Cells in droplets after culturing for 3.5 days



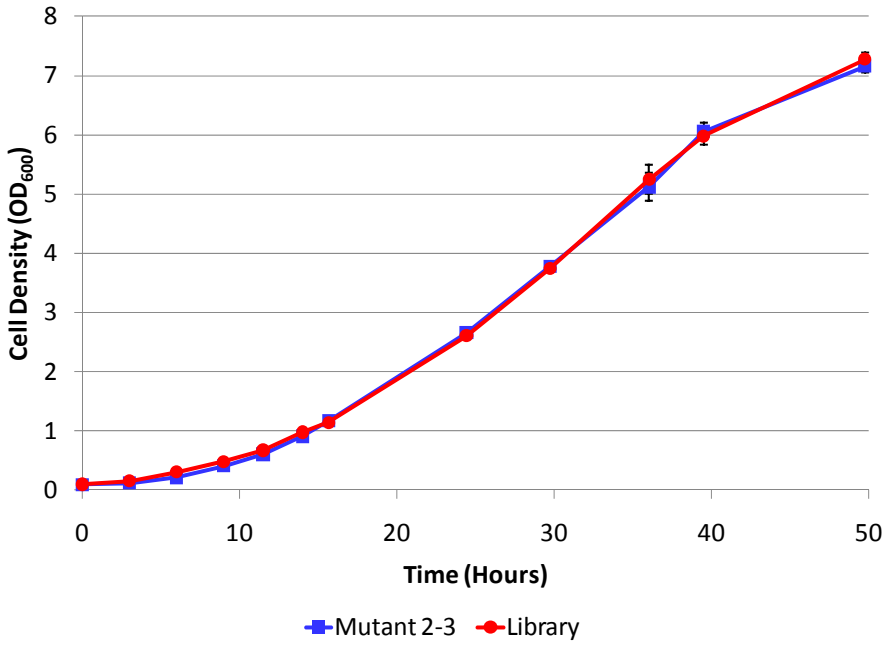
Supplementary Figure 8 Mask design for second device where reinjected droplets and assay droplets are coalesced to initiate the assay reaction, droplets flow through the delay lines allow the assay reaction to proceed, and the droplets are sorted based on the fluorescence intensity. Desired droplets flow into the upper sorting channel and undesired droplets flow into the lower channel.



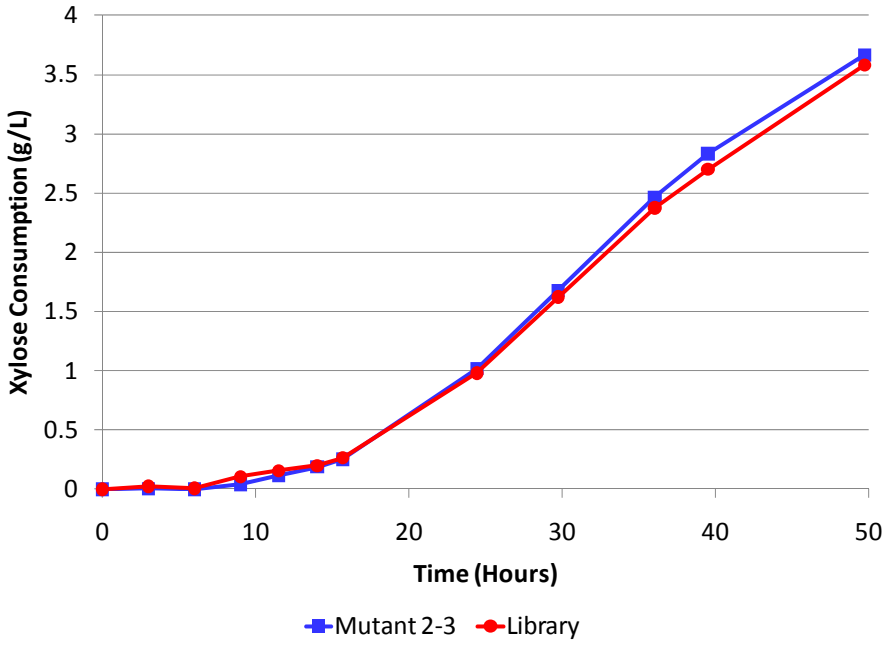
Supplementary Figure 9 Microfluidic device to create calibration curves. It contains one oil input and two aqueous inputs. Fluorescence values reported in Supplementary Figure **10** were measured in the middle of the 10th delay line.



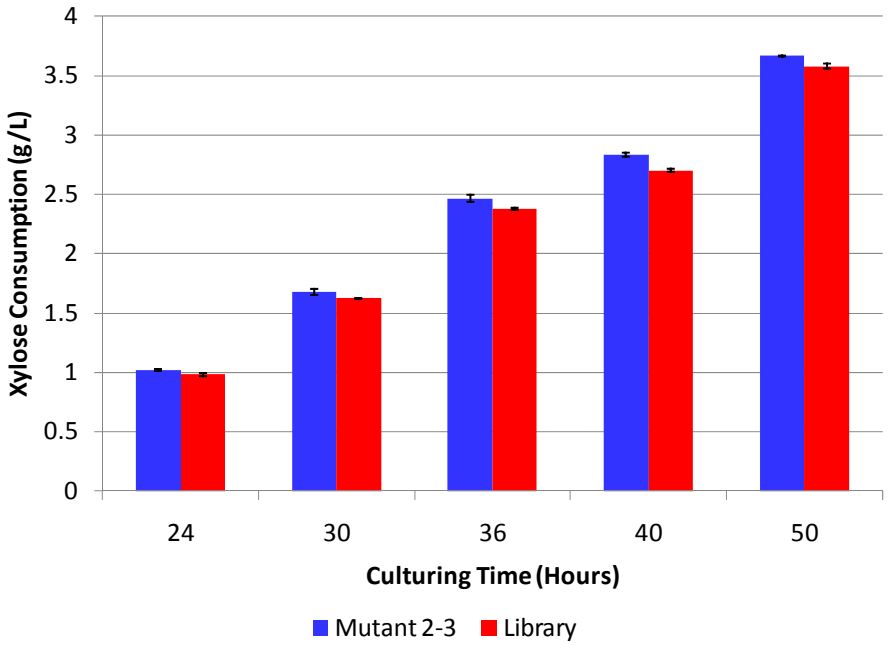
Supplementary Figure 10 Fluorescence vs. xylose calibration curve for four concentrations (0, 1, 2.5, and 5g/L) of xylose. The xylose concentration supplied to the device is reported in this graph



Supplementary Figure 11 Cell growth comparison of mutant 2-3 and the library as a whole



Supplementary Figure 12 Xylose consumption comparison of mutant 2-3 and the library as a whole



Supplementary Figure 13 Xylose consumption comparison of mutant 2-3 and the library as a whole during mid to late exponential phase. At the 40 hour time point, mutant 2-3 had a 4.9% higher xylose consumption and the difference was statistically different ($p < 0.05$ using a two-sided t-test). Error bars denote the standard deviation from biological replicate experiments ($n = 2$).

Supplementary Table

Experiment	False Positive Error Rate
1:1 H131/TAL1 Enrichment (Bin < 0.6)	2.4%
1:1 H131/TAL1 Enrichment (Bin < 0.7)	2.6%

Supplementary Table 1: False positive error rate from enrichment experiments

MOVIES

Supplementary Movie 1 Droplet formation of yeast or *E. coli* cells in growth medium

Supplementary Movie 2 Alternating sequence of cell-containing droplets and assay droplets

Supplementary Movie 3 Droplet coalescence of cell-containing droplets and assay droplets

Supplementary Movie 4 Sorting of "desired" droplets into the upper channel

REFERENCES

1. Ahn, K., Agresti, J., Chong, H., Marquez, M. & Weitz, D.A. Electrocoalescence of drops synchronized by size-dependent flow in microfluidic channels. *Applied Physics Letters* **88**, 264105 (2006).
2. Huebner, A. et al. Quantitative detection of protein expression in single cells using droplet microfluidics. *Chem. Commun.* **12**, 1218-1220 (2007).
3. Ahn, K. et al. Dielectrophoretic manipulation of drops for high-speed microfluidic sorting devices. *Applied Physics Letters* **88**, 024104 (2006).
4. Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403-410 (1990).
5. Saloheimo, A. et al. Xylose transport studies with xylose-utilizing Saccharomyces cerevisiae strains expressing heterologous and homologous permeases. *Applied Microbiology and Biotechnology* **74**, 1041-1052 (2007).