How to cite: *Angew. Chem. Int. Ed.* **2023,** *62,* e202303112
doi.org/10.1002/anie.202303112

Enzymes Hot Paper

Controlled Continuous Evolution of Enzymatic Activity Screened at Ultrahigh Throughput Using Drop-Based Microfluidics

R. G. Rosenthal[+], X. Diana Zhang[+], K. Ilić Đurđić, J. J. Collins, and D. A. Weitz**

Abstract: Enzymes are highly specific catalysts delivering improved drugs and greener industrial processes. Naturally occurring enzymes must typically be optimized which is often accomplished through directed evolution; however, this is still a labor- and capital-intensive process, due in part to multiple molecular biology steps including DNA extraction, in vitro library generation, transformation, and limited screening throughput. We present an effective and broadly applicable continuous evolution platform that enables controlled exploration of fitness landscape to evolve enzymes at ultrahigh throughput based on direct measurement of enzymatic activity. This drop-based microfluidics platform cycles cells between growth and mutagenesis followed by screening with minimal human intervention, relying on the nCas9 chimera with mutagenesis polymerase to produce in vivo gene diversification using sgRNAs tiled along the gene. We evolve alditol oxidase to change its substrate specificity towards glycerol, turning a waste product into a valuable feedstock. We identify a variant with a 10.5-fold catalytic efficiency.

Introduction

Enzyme-catalyzed transformations are applied in a diverse range of fields, including drug synthesis, household products and research and industrial processes.[1] Enzymes can be designed to convert waste products into valuable materials or help degrade waste that is not otherwise biodegradable, making them very environmentally friendly. Enzymatic catalysis is highly specific and efficient yet can operate with very high turnover while still under mild conditions without using toxic materials of any sort. Typically, enzymes found in nature require adaptations to function optimally in a specific application. Enzyme activity can be optimized through fully guided methods like rational computational design or through trial-and-error-based methods like directed evolution.[1a] Despite the great strides made in rational design-based methods, they are still too complex to be applicable in many cases. Instead, directed evolution remains a powerful and widely used technique to optimize enzymes for tailored applications. It entails generating genetic diversity and selection of variants with desired phenotype.[2] Typical directed evolution Schemes comprise multiple rounds, with each round requiring several steps including DNA extraction, in vitro library generation, transformation, and screening.[3] Such screening for enzyme variants of higher activity can be done at a very high throughput using droplet microfluidics, where the enzymatic activity of a single bacterial or yeast cell can be measured at throughputs of more than 10^7 variants per day, yielding significantly improved variants.[4]

Despite the success of directed evolution, it is, however, still a labor- and capital-intensive process.[1c] Converting this labor-intensive, stepwise process of directed evolution into a more streamlined and continuous process would make it much more broadly applicable. To achieve such continuous evolution, both gene diversification and selection or screening must occur in an autonomous fashion.[3b] One method to select for enzymatic activity with minimal user input is by creating conditions under which the target enzyme catalyzes

[*] R. G. Rosenthal,[+] X. Diana Zhang,[+] K. I. Đurđić, D. A. Weitz
School of Engineering and Applied Sciences, Harvard University
Cambridge, MA 02138 (USA)
E-mail: kilicdurdic@seas.harvard.edu
weitz@seas.harvard.edu

R. G. Rosenthal,[+] J. J. Collins
Institute for Medical Engineering and Science, Massachusetts Institute of Technology
77 Massachusetts Ave, Cambridge, MA 02139 (USA)

K. I. Đurđić, J. J. Collins, D. A. Weitz
Wyss Institute for Biologically Inspired Engineering, Harvard University
3 Blackfan Circle, Boston, MA 02115 (USA)

K. I. Đurđić
University of Belgrade-Faculty of Chemistry
Studentski trg 12–16, 11000 Belgrade (Serbia)

J. J. Collins
Department of Biological Engineering, Massachusetts Institute of Technology
77 Massachusetts Ave, Cambridge, MA 02139 (USA)
Synthetic Biology Center, Massachusetts Institute of Technology
77 Massachusetts Ave, Cambridge, MA 02139 (USA)
Harvard-MIT Program in Health Sciences and Technology, Massachusetts Institute of Technology
77 Massachusetts Ave, Cambridge, MA 02139 (USA)
Broad Institute of MIT and Harvard
415 Main St, Cambridge, MA 02142 (USA)

D. A. Weitz
Department of Physics, Harvard University
Cambridge, MA 02138 (USA)

[+] These authors contributed equally to this work.

a rate-limiting conversion for cellular growth.[5] Growth-based selection makes faster growth a proxy for higher enzymatic activity. While this approach enables high screening throughput, it often requires intricate genetic designs to couple cellular or viral growth to the trait to be evolved. Instead, directly assaying enzymatic activity is desired in most cases.[6] Many approaches exist to directly assay enzymatic activity; however, they often rely on cell lysis and measuring enzymatic activity in lysate or purified protein. Cell lysis, however, is incompatible with a continuous process. To overcome this limitation, the target enzyme must be either secreted or displayed on the cell surface.[7]

Another crucial part of directed evolution is gene diversification. Ideally, only the target sequence is mutated, minimizing the chance that off-target mutations give rise to the selected phenotype without improving the desired phenotype. Off-target mutations in the promoter region can, for example, increase the expression levels instead of the enzymatic activity.[8] Such mutations give rise to increased enzymatic turnover, which is the selected phenotype, without improving enzymatic activity, which is the desired phenotype. These effects can lead to false positives and false negatives in the selection process. In addition, introducing targeted mutations increases the throughput of screening schemes by orders of magnitude because no screening capacity is wasted on off-target mutations. In vivo gene diversification can generate genetic diversity without the need for ex vivo molecular biology steps; instead, it relies on DNA modifying enzymes expressed in the cell. It has been shown that nCas9 can be used as a targeted gene diversification tool, however, it has not been applied to evolving traits, such as enzymatic activity, unrelated to cell growth.[3b,9] Development of a method for continuous directed evolution which does not require complex molecular biology and can screen directly for enzymatic activity at high throughput would greatly simplify directed evolution workflows and significantly improve our ability to tailor enzyme activity for specific applications.

In this paper, we develop a continuous evolution platform to evolve enzymes at ultrahigh throughput based on a direct measurement of enzymatic activity. We study the sugar oxidizing enzyme, alditol oxidase from *Streptomyces coelicolor*, and change its substrate specificity towards glycerol, changing a waste product into glycerate, a valuable feedstock.[10] Genetic diversity is generated with the combination of improved versions of the nCas9 based EvolvR in vivo gene diversification system and single-guide RNAs (sgRNAs) tiled along the full length of the gene.[9] From the generated genetic diversity, we sort the cells expressing variants with the highest enzymatic activity each round. We use a single microfluidic chip that mixes cells with reagents, incubates the cells for the duration of the reaction, and sorts the best variants. This system enables continuous passage of cells between single-cell in-droplet screening, and growth and diversification, with minimal intervention. We achieve a screening throughput of $>2\times10^6$ variants per day and can run the continuous evolution over multiple days, providing an efficient means of identifying improved variants with a single mutation. Seeding the continuous evolution platform with hits identified in this initial screen yields multiple variants with significantly improved activity through additional mutations. We exploit this system to isolate rare mutations which greatly improve the oxidation of glycerol with molecular oxygen catalyzed by alditol oxidase.

Results and Discussion

Sugar oxidizing enzymes are used in applications ranging from drug delivery to blood glucose detection.[11] Many applications require tailoring the substrate specificity to the specific application. Here we develop and apply an efficient continuous evolution platform to change the substrate specificity of alditol oxidase from *Streptomyces coelicolor* towards glycerol, producing an alternative to the chemical oxidation, turning a waste product into a valuable feedstock, glycerate.[10] To develop an effective and efficient continuous evolution platform, we combine a system that introduces mutations in a targeted fashion into the gene of interest in vivo with a comprehensive microfluidic device that encapsulates single cells in droplets, adds substrate, incubates for a controlled period, and sorts cells, based on enzymatic activity, directly into fresh growth medium enabling further diversification. To generate genetic diversity, we introduce a mutagenic polymerase. To introduce mutations more effectively, we use a fusion of a nicking CAS9 (nCas9) to a mutagenic polymerase (nCas9-PolI) enabling us to target the mutations through the introduction of a single guide RNA (sgNRA), see Figure 1A.[9] To assay enzymatic activity, we use an ice-nucleation-protein (INP) based enzyme display system, from *Pseudomonas syringae*, which is transported to the outside of the cell, allowing contact with the substrate, see Figure 1B.[7,12] Enzymatic activity is determined from the fluorescence level detected using a photomultiplier tube, enabling screening of drops above a threshold fluorescence level. To operate the continuous evolution, we use a lab-on-a-chip system that automates cell encapsulation, incubation, and screening.

Generating genetic diversity with nCas9-PolI requires adding a sgRNA to the plasmid to target the mutations to a specific DNA sequence. To initialize diversification, plasmids containing the nCas9-PolI with a sgRNA and the target gene are co-transformed into *E. coli*. With both the mutagenesis system and target gene present, mutations accumulate in the target gene during growth. To measure the mutation rate under different conditions, we use a green fluorescent protein (GFP) with an early stop codon, and a sgRNA targeting the nCas9-PolI to the stop codon. The mutation rate can be estimated by measuring the fraction of cells that gain fluorescence with a flow cytometer 24 hours after co-transforming the GFP and mutagenesis constructs. With the initial mutagenesis conditions, about 0.03 % of cells contain a mutation after 24 hours.[9] To avoid screening unmutated sequences, we optimize the mutagenesis system by varying the temperature since lower temperatures increase the binding affinity of DNA polymerase I to DNA.[13] Therefore, we reason that a lower temperature might have a positive impact on the window of mutagenesis

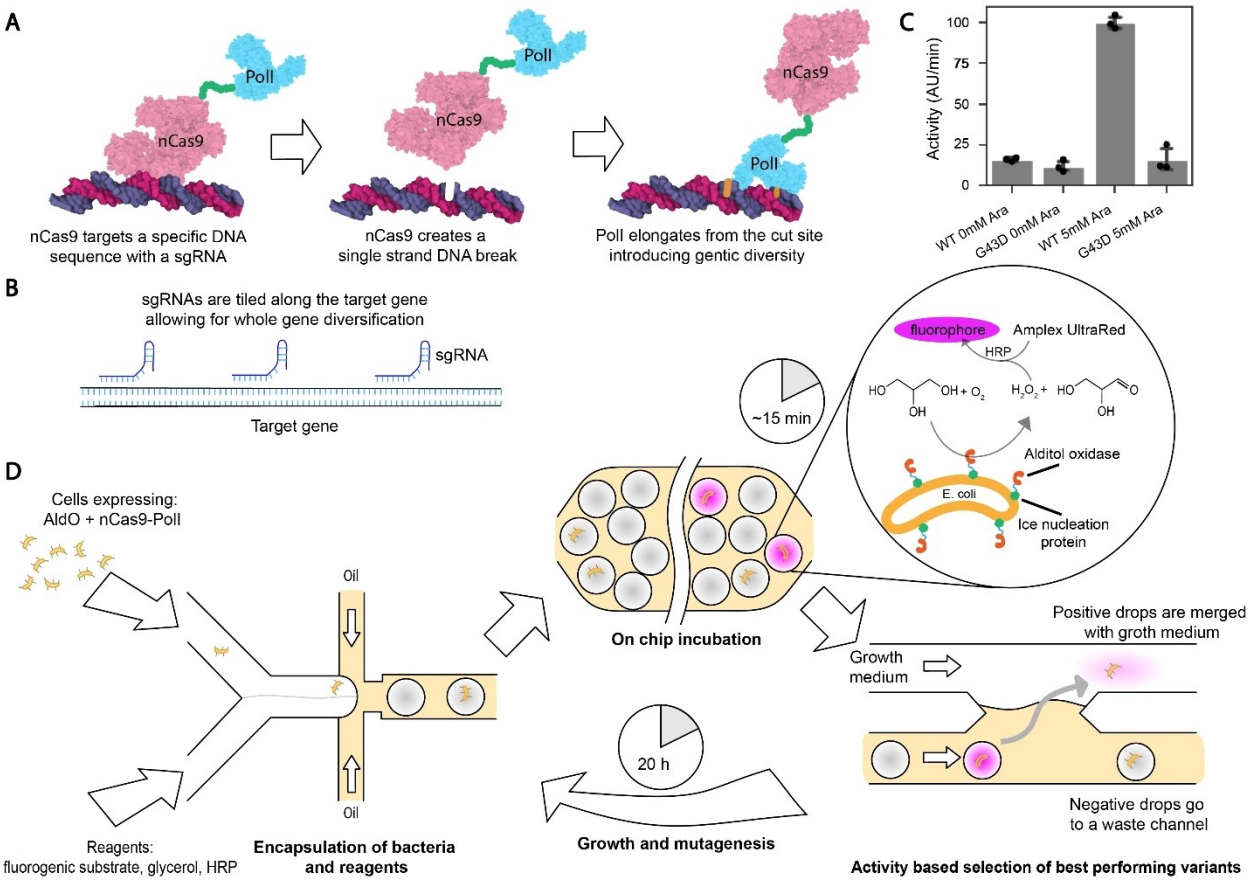

Figure 1. Activity-based continuous evolution platform. A) Overview of the in vivo genetic diversification construct. A nicking CAS9 makes a single-strand break in the template DNA and a mutagenic PolI extends the nicked strand, generating genetic diversity in the downstream sequence. B) Multiple constructs with tiled sgRNAs are transformed to ensure mutations occur along the target gene. C) AldO tethered to the cell's surface with ice nucleation protein is active and only expressed in the presence of the inducer, arabinose. AldO G43D is an inactivated mutant. D) Schematic overview of the continuous evolution platform. Single cells are encapsulated in ≈5 pL aqueous drops containing assay reagents, followed by an on-chip incubation for about 15 minutes. After the incubation step, drops containing cells with active alditol oxidase will be fluorescent. The AldO oxidizes glycerol with molecular oxygen, forming glyceraldehyde and H_2O_2, which is a co-substrate in the horseradish catalyzed oxidation of Amplex UltraRed™ into a fluorescent product. The fluorescent signal of each drop is proportional to the amount of glycerol oxidized by the AldO displayed on the cell's surface. Drops with a fluorescent signal above the set threshold, are merged with a stream of growth medium and the recovered cells are grown for ≈20 h, before being subjected to another round of selection.

downstream of the sgRNA binding site. Consistent with our expectations, reducing the growth temperature from 37 °C to 28 °C increases the fraction of fluorescent cells after 24 hours more than 30-fold, from 0.03 to roughly 1 % (Figure S1). Initial experiments show that inducing the expression of nCas9-PolI with anhydrotetracycline led to lower mutation rates, suggesting that lower expression could improve mutation rates. Thus, to further optimize the system, we lower the expression of nCas9-PolI by using different promotors (Table S1) that affect the mutation rate, the mutation window, and the growth speed to varying degrees.[14] The doubling time of *E. coli* containing different promoter constructs varies from 1.3 hours to 2.8 hours at 28 °C, while the mutation rate, as measured by the fraction of GFP positive cells, varies from 0.0 to 6.3 %, with no apparent correlation between growth and mutation rates (Figure S2–6). We, therefore, use all promoters and PolI constructs which show a GFP positive cell fraction of >1% after 24 h.

To screen beneficial mutations in AldO, we establish a fluorescence-based activity assay. In the assay, AldO is expressed tethered to *E. coli* by an INP and converts the substrate glycerol with molecular oxygen to glycerate and hydrogen peroxide. The hydrogen peroxide that is generated is then converted to a fluorescence signal by the horseradish peroxidase mediated conversion of Amplex™ UltraRed into a fluorescent product. We measure activity with and without expression inducer, arabinose, for both wild-type AldO and inactivated variant, AldO G43D. We detect increase in fluorescent signal only for the wild-type AldO in presence of inducer, confirming that signal is a result of enzyme activity (Figure 1C). The accumulated fluorescent product is quantified on a single-cell level by a custom microscope setup with a 543 nm excitation laser thereby stratifying cells based on their enzymatic activity (Figure 1D).

The assay is used in a custom lab-on-a-chip microfluidic device that is operated under steady-state conditions to screen for the best performing AldO variants. During the operation of the device, drop making, incubation, and sorting occur at the same frequency, ensuring consistency in the incubation time and allowing the chip to operate continuously. At the drop-making junction, AldO-expressing cells are mixed with assay regents and encapsulated in monodisperse droplets that are about 5 pL in volume (Figure S7). The drops produced flow into the incubation line where the reaction takes place, generating a fluorescence signal proportional to the enzymatic activity. The sorting module at the end of the incubation line merges those drops that are above a fluorescence threshold directly with PBS solution in an adjacent channel, allowing the system to be operated continuously. The system is operated at a cell loading of around 1 cell per 12 drops, reducing the frequency of double-loaded drops to below 1 per 300 drops. The chip operates at 1500–1800 drops s^{-1}, thus probing $\approx 5 \times 10^5$ cells h^{-1}. The measured sorting enrichment is $>98\,\%$. A typical experiment is run for 3 to 4 hours and screens about 2×10^6 cells. Screening at the single-cell level leads to a weaker genotype-to-phenotype linkage because the cells are at different growth stages and stochastic noise in gene expression is not averaged over a population. We overcome this limitation by screening at ultrahigh throughputs, effectively screening the same enzyme variant across multiple cells. More importantly, we run our continuous evolution platform for multiple days; this increases the diversity while also averaging stochastic noise over sequential rounds of screening.

To ensure that mutations occur along the whole AldO gene, we design eight sgRNAs with high specificity scores[15] which are tiled along the AldO gene, about 150 base pairs apart, as shown in Figure 1B. With this layout, the mutation windows are spaced roughly evenly along the whole length of the AldO gene. Continuous evolution is initiated by co-transforming a pair of plasmids, the first containing the nCas9-PolI construct together with one of the targeting sgRNAs and the second containing the AldO target plasmid. The cells are first grown for 24 hours to allow gene diversification to occur and then screened. The top $\approx 0.5\,\%$ of best-performing cells are sorted, resulting in about 500–5000 isolated cells (Figure 2A). The selection criterion at this stage is not very stringent to ensure that many mutants are selected. These selected cells are grown overnight and subjected to another round of screening the next day where the top 0.5 % are again selected. At this point, many copies of each of the mutants selected in the first round is screened, thereby increasing their contribution to newly selected cells, while, at the same time, new mutations from the second round are also selected. This process is repeated for four days which further enriches the most active mutants while also selecting newly mutated variants. After the four rounds the cells are re-transformed to prevent the accumulation of off-target mutations that lead to excessively slow cell growth. The whole process of continuous diversification and selection followed by retransformation is repeated two more times for three days each, for a total of 10 rounds. To validate the progress of the continuous evolution, a population of evolved AldO variants after five days is compared to AldO WT which is co-transformed with the mutagenesis plasmid and grown for one day without being subjected to screening. When passed through the continuous evolution chip, the evolved population displays an extended tail at the higher end of the activity distribution, compared to the unevolved population as shown in the histogram of detected fluorescence intensities plotted in Figure 2A. However, the mean activity of the cell-containing drops stays unchanged, underlining the importance of multiple selection rounds to enrich the most active variants. On the 3rd, 5th, 7th, and 9th evolution day, the plasmid DNA is isolated and the AldO plasmids are re-transformed into cells in the absence of the mutagenesis plasmid and plated onto agar. Single colonies are transferred to separate wells of a 96-well plate and clonally expanded to remove stochastic

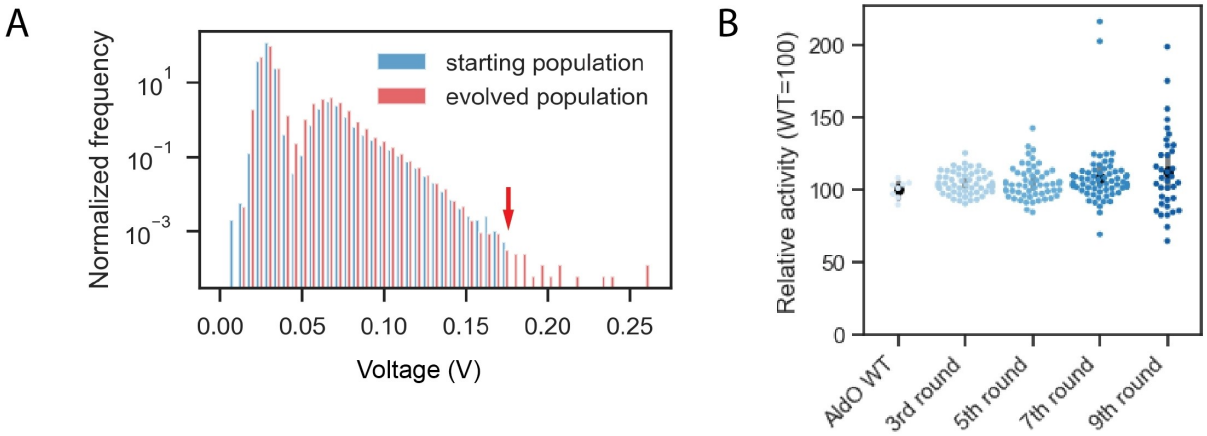

Figure 2. Activity-based continuous evolution of Alditol oxidase. A) A representative comparison of the activity distribution of the starting population and the activity distribution after 5 days of continuous evolution. The evolved population has many more cells at the high end of the activity distribution. B) Validated activities of cell samples along the continuous evolution trajectory, displaying the enrichment of variants that have increased activities compared to WT AldO.

noise associated with single-cell assays. A fluorescence-based validation assay is performed using a plate reader to measure activity. The AldO variants taken further along the continuous evolution trajectory, display an extended tail at the higher end of the activity distribution and a larger variance (Figure 2B), similar to the results of the assay performed in the continuous evolution chip.

To confirm the efficiency of our continuous evolution process, we sequence plasmids after the 3rd, 5th, 7th, and 9th days of evolution using a Pacbio sequencer to obtain long reads. We restrict our analysis to highly accurate plasmid consensus sequences with the correct AldO gene length and focus only on nucleotide substitutions. The number of WT nucleotide sequences decreases dramatically by the 3rd day and even more by the 5th day; thereafter, there is a slow but steady decrease until the 9th day of continuous evolution as shown by the blue bars in Figure 3A. However, single nucleotide mutations can also lead to silent contributions to the amino acid sequence, and thus the number of WT amino acid sequences decreases slightly after the 3rd day but then actually begins to increase by the 9th day. Accumulation of silent mutations in WT sequence may result in higher expression levels. Also, buy this time in selection process a big portion of sequences are already mutated and the mutational burden is increasing, resulting in a higher number of inactive mutants. Nevertheless, the result is an increase in the fraction of WT amino acid sequences, but no WT DNA sequencies. By comparison, the number of AldO variants with different amino acid sequences per 1000 analyzed full gene length sequences increases dramatically on the 3rd day but then decreases significantly on the subsequent days, as shown in Figure 3B. This reflects the initial increase in mutations followed by the sorting of the best variants. By the 9th day, only eight variants remain per 1000 sequences analyzed, reflecting the efficiency of the ultrahigh-throughout droplet microfluidics-based screening system in sorting the most active variants. Analysis of mutation distribution after 3rd day of evolution also shows that sgRNA guides are positioned such that mutations are introduced throughout the entire gene sequence, with slightly lower frequency in the beginning and the end of the gene. However, this slight bias is very well addressed with microfluidics-based high throughput screening (Figure S7).

To follow the continuous evolution of amino acid sequences, we compile a table whose columns list the top 14 mutated amino acid positions, each of which is greater than 1 % of the total at some point. The rows of the table list the mutation in each position after the 3rd, 5th, 7th, and 9th day of

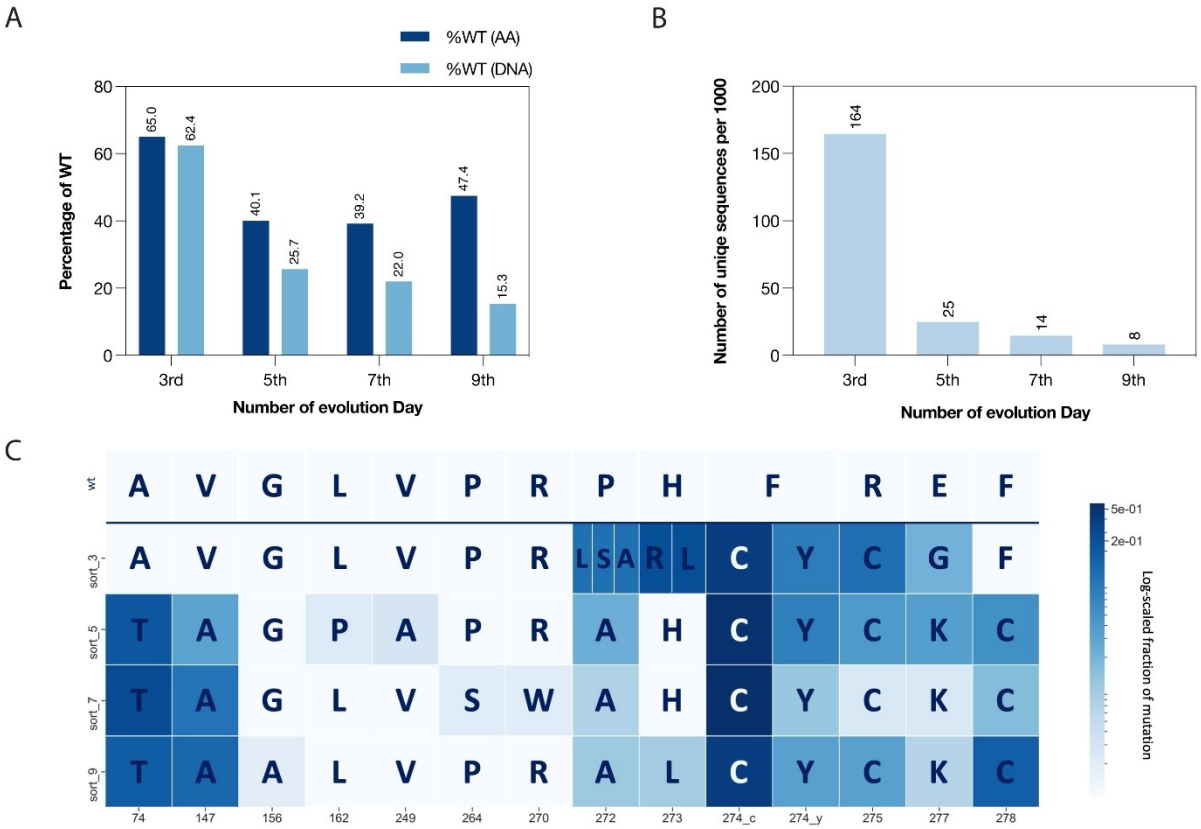

Figure 3. Full gene length sequencing analysis of AldO through continuous evolution process. A) Representation of percentage of WT nucleotide and amino acid sequences through the continuous evolution process. B) Representation of the decrease in the number of different variants, normalized by 1000 analyzed sequences, through the continuous evolution process. C) Heatmap representation of the most abundant mutations (appearing as at least 1 % of total analyzed sequences) through the continuous evolution process. The scale from white to dark blue indicates an increase in the number of gene copies containing that substitution.

evolution and screening. The color map reflects the abundance of the mutation with darker colors representing more mutations. Beneficial mutations appear immediately upon the first diversification step, followed by their enrichment during the subsequent screening and their substitution with more active variants later in the process, as shown in Figure 3C. One of the most abundant mutations in the early stages of continuous evolution is F274Y which is represented as 2.50 % of all sequences after the 3rd day and is enriched to 4.92 % after the 5th day; however, the appearance of more beneficial mutations ultimately reduces the representation of F274Y to 0.46 % after the 7th day. Another variant, F278 C, is only 0.04 % of total analyzed sequences after the 3rd day but is enriched to 11.41 % after the 9th day. This same enrichment process is observed for another variant, V147 A, which first appears after the 5th day and rises to 9.59 % of all analyzed sequences after the 9th day.

To further characterize the most enriched variants we express them, as well as WT AldO and previously reported variant V125M V133M A244T G399R (M1) (*17*) in soluble form and purify them using a NiNTA matrix (Figure S12). We measure steady-state kinetic parameters for glycerol as a substrate using ABTS-HRP coupled assay. The most significant improvement in both catalytic parameters is observed for F278 C variant where 1.8-fold improvement in k_{cat} and 5.8-fold decrease in K_M value result in 10.5-fold higher specificity constant for glycerol as a substrate when compared with WT enzyme. We report two additional variants with improved specificity for a glycerol as a substrate: A74T and V147 A. These variants exhibit 2.55 and 1.96-fold improvement in specificity constants (Table 1). Two significantly enriched variants: F274 C and A74T F274 C could not be expressed in soluble form. This presumably reflects the difference between the surface displayed state, which is the basis for the selection and the intracellular expression.

In this study, we develop and implement a continuous evolution platform that enables controlled exploration of fitness landscape to evolve enzymes at ultrahigh throughput based on direct measurement of enzymatic activity. The mutation rate of the continuous evolution process is very low compared to the more commonly used error prone PCR workflow; it is still 5–6 times higher than background level. We set the mutational burden of the dCAS9-PolI construct at 10^{-4} mutations base pair^{-1} generation^{-1}. However, because the process is continuous, the cells undergo many divisions thereby increasing the overall mutational burden considerably. Moreover, each mutation is most likely a single mutation, which has the highest probability of yielding a beneficial result. Studies on TEM-1 beta-lactamase estimate that 7 % of single mutations are beneficial whereas only 0.07 % of double mutations are beneficial.[16] By comparison, mutations generated by error-prone processes follow Poisson statistics and only the average mutation rate can be influenced experimentally.[17] To ensure that the majority of mutations are single mutations yielding the highest likelihood of beneficial results requires a low average mutations rate and hence most genes are unmutated. If instead the mutational burden is increased, most variants have multiple mutations and are mostly less active or completely inactive compared to the starting point. Thus, the continuous evolution process, is in fact, an efficient way of generating beneficial variants consisting of a single mutation. By converting the directed evolution process from stepwise to continuous, we can easily cycle cells many times between diversification and screening without the necessity of developing selection-based assays. The many cycles allow for a gradual exploration of the fitness landscape, albeit only in the vicinity of the seeded sequence(s).

As the continuous evolution proceeds the population is increasingly shifted to improved variants raising the average activity level; as a result only variants with increasing activity will be selected as the evolution proceeds. For example, one variant that appears very early in continuous evolution F274Y initially becomes more prevalent in first few rounds, but ultimately does not persist as the evolution proceeds. This is because catalytic efficiency is not significantly greater than the WT are ultimately reduced as variants that are significantly improved, such as F278C, begin to dominate the population. Thus, this leads to an efficient way of exploring the space of variants with single mutations.

To explore a larger region of the fitness landscape and to obtain even better performance requires variants with multiple mutations. The continuous evolution process is more efficient in exploring variants with a single mutation, although some variants with two mutations do appear, such as A74T F274C. To increase the likelihood of obtaining variants with the second mutation using continuous evolution, we use a purifying selection to remove all contributions from the WT to the screen, thereby eliminating the silent mutations that become dominant. Instead, we choose four initial variants that each have one mutation, enabling the dCAS9-PolI to add a second mutation in a controlled manner. After three rounds following the purifying selection, we do find a variant the has an additional mutation, but its activity is only improved when is displayed on the cell

Table 1: Steady-state kinetics parameters for AldO WT and selected variants.

Variant	k_{cat} [10^3 s^{-1}]	K_M [mM]	k_{cat}/K_M [10^4 $s^{-1}M^{-1}$]	Improvement [fold]
WT	4.29	288.3	1.49	1
F278C	7.81	49.95	15.64	10.5
A74T	6.80	178.7	3.81	2.55
V147A	6.41	219.7	2.92	1.96
F274Y	3.90	266.3	1.44	0.97
V125M, V133M, A244T, G399R (M1)	2.24	72.91	3.073	2.1

surface, but not when it is expressed as a soluble protein. Additional rounds of continuous evolution may lead to variants with improved activity that can be translated to a soluble enzyme.

Nevertheless, this continuous evolution process provides an efficient method to explore the fitness landscape in a controlled manner.[18] Moreover, this continuous evolution method with our microfluidics platform enables a very large number of cells to be screened without the need for in vitro molecular biology steps. We are able to identify a variant that is 10.5 times more catalytically efficient than the WT.

The best variant we find is AldO F278 C. By cycling cells between screening, growth, and mutagenesis, we mimic natural evolution but at a pace that is orders of magnitude faster, yielding a 10.5-fold improved AldO variant within days. By contrast, screening for the same improvement in glycerol oxidation rate by AldO using traditional multi-step error-prone PCR and combinatorial mutagenesis resulted in only a 2.4-fold improvement (Table 1).[10] These results highlight the power of the ultrahigh-throughput continuous evolution approach introduced here to identify beneficial mutations. Importantly, the result of our study is not just a more active AldO variant, but a highly efficient *E. coli*-based whole-cell biocatalyst suitable for direct biotechnological applications.[7,12]

In this work, we demonstrate the power of activity-based, continuous evolution to improve enzymatic activity towards alternative substrates. Using current setup, accumulation of off-target mutations in *E. coli* cells over the time leads to slower cell growth. We overcome this issue by isolating plasmids 3 times during evolution process and transforming fresh cells. This could also be addressed by using engineered *E. coli* cells with reduced and stabilized genome and removed genes for stress-induced diversity generated DNA polymerases, resulting in lower mutation rates.[19] It would also be of interest to investigate how quickly different enzyme variants can be adapted to new selection pressures. In addition, to achieve further improvements of the variants, the purifying selection could be applied multiple times, until the desired activity or a physical constraint is reached. In an industrial setting, the continuous evolution method could be seeded with multiple enzyme variants which occur in nature as input. By choosing more diverse input sequences, no variant will be too dominant in the higher end of the activity distribution. The continuous evolution platform enables these questions to be addressed at a throughput not possible thus far. In addition, applying eight separate sgRNAs allows targeting of the full AldO gene; however, it would also be possible to use our strategy to identify selected regions of the gene sequence that are important for increased activity toward the desired substrate and start a new continuous evolution process with only sgRNAs targeting these regions. Also, to achieve higher mutation rates with multiple mutations throughout the whole gene without the need for re-transformation of cells with new sgRNA all sgRNA-coding genes could be expressed from the same plasmid in the same cell. Although, the nCas9-EvolvR in vivo gene diversification system can be uncoupled from microfluidics, low mutation rates in absence of a high throughput screening platform would result mostly in measurements of WT enzyme variants or would require development of complex selection-based assays where the actual activity of the enzyme cannot be measured but only assessed as a function of cell survival or death.

Conclusion

We demonstrate a novel continuous evolution platform for improvement of enzymatic activity that functions at ultra-high throughput. Using a combination of cell surface display, nCas9-mutagenic polymerase and drop microfluidics-based cell sorting we change specificity of alditol oxidase towards glycerol, converting a waste product into a valuable feedstock. Demonstrated platform enables controlled exploration of enzyme fitness landscape and results in a variant with over 10-fold higher than WT catalytic efficiency after only 10 days of evolution, with minimal human input. Developed methodology enables both full gene-length and targeted mutagenesis and is easily translatable to many industrially important enzymes.

Author Contributions

Conceptualization: RGR, JJC, and DAW; Investigation: RGR and XDZ; Data analysis, Validation: RGR, XDZ and KIĐ; Writing-original draft, Writing-review & editing: RGR, XDZ, KIĐ, and DAW.

Data and materials availability

All data, code and materials used in this study are available from corresponding authors.

Acknowledgements

We thank Gael Blivet and Sofiane Decombas for their help in early stages of the project. Funding: This work was supported in part by the National Science Foundation (CBET-2103538), the Harvard MRSEC (DMR-2011754), the Wyss Institute and Google. R.G.R. was supported by an EMBO Fellowship (ALTF 1041-2016) and a Branco Weiss Fellowship.

Conflict of Interest

The authors declare no competing interest.

Data Availability Statement

The data that support the findings of this study are available in the supplementary material of this article.

Keywords: CRISPR-Directed Evolution · Continuous Evolution · Droplet Microfluidics · Protein Engineering · Sustainable Chemistry

[1] a) S. Wu, R. Snajdrova, J. C. Moore, K. Baldenius, U. T. Bornscheuer, *Angew. Chem. Int. Ed.* **2021**, *60*, 88–119; b) J. Chapman, A. E. Ismail, C. Z. Dinu, *Catalysts* **2018**, *8*, 238; c) I. Victorino da Silva Amatto, N. Gonsales da Rosa-Garzon, F. Antônio de Oliveira Simões, F. Santiago, N. Pereira da Silva Leite, J. Raspante Martins, H. Cabral, *Biotechnol. Appl. Biochem.* **2022**, *69*, 389–409.

[2] a) F. H. Arnold, *Angew. Chem. Int. Ed.* **2018**, *57*, 4143–4148; b) S. Gargiulo, P. Soumillion, *Curr. Opin. Chem. Biol.* **2021**, *61*, 107–113.

[3] a) R. Martínez, U. Schwaneberg, *Biol. Res.* **2013**, *46*, 395–405; b) B. Álvarez, M. Mencía, V. de Lorenzo, L. Á. Fernández, *Nat. Commun.* **2020**, *11*, 6436.

[4] a) J. J. Agresti, E. Antipov, A. R. Abate, K. Ahn, A. C. Rowat, J. C. Baret, M. Marquez, A. M. Klibanov, A. D. Griffiths, D. A. Weitz, *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 4004–4009; b) R. Prodanović, W. L. Ung, K. Ilić Đurđić, R. Fischer, D. A. Weitz, R. Ostafe, *Molecules* **2020**, *25*, 2418; c) A. Stucki, J. Vallapurackal, T. R. Ward, P. S. Dittrich, *Angew. Chem. Int. Ed.* **2021**, *60*, 24368–24387.

[5] S. Bershtein, M. Segal, R. Bekerman, N. Tokuriki, D. S. Tawfik, *Nature* **2006**, *444*, 929–932.

[6] H. Xiao, Z. Bao, H. Zhao, *Ind. Eng. Chem. Res.* **2015**, *54*, 4011–4020.

[7] L. Han, Y. Zhao, S. Cui, B. Liang, *Appl. Biochem. Biotechnol.* **2018**, *185*, 396–418.

[8] a) D. M. Weinreich, N. F. Delaney, M. A. DePristo, D. L. Hartl, *Science* **2006**, *312*, 111–114; b) G. Muteeb, R. Sen, *Methods Mol. Biol.* **2010**, *634*, 411–419.

[9] S. O. Halperin, C. J. Tou, E. B. Wong, C. Modavi, D. V. Schaffer, J. E. Dueber, *Nature* **2018**, *560*, 248–252.

[10] S. Gerstenbruch, H. Wulf, N. Mußmann, T. O'Connell, K.-H. Maurer, U. Bornscheuer, *Appl. Microbiol. Biotechnol.* **2012**, *96*, 1243–1252.

[11] a) P. Mousavi, B. Khalvati, A. Maleksabet, A. Savardashtaki, M. Taheri-Anganeh, A. Movahedpour, *Biotechnol. Appl. Biochem.* **2021**, *69*, 939–950; b) L. Zhao, L. Wang, Y. Zhang, S. Xiao, F. Bi, J. Zhao, G. Gai, J. Ding, *Polymers* **2017**, *9*, 255.

[12] E. van Bloois, R. T. Winter, D. B. Janssen, M. W. Fraaije, *Appl Microbiol Biotechnol* **2009**, *83*, 679–687.

[13] K. Datta, A. J. Wowor, A. J. Richard, V. J. LiCata, *Biophys. J.* **2006**, *90*, 1739–1751.

[14] B. Wang, M. Barahona, M. Buck, *Nucleic Acids Res.* **2015**, *43*, 1955–1964.

[15] G. Liu, Y. Zhang, T. Zhang, *Comput. Struct. Biotechnol. J.* **2020**, *18*, 35–44.

[16] C. E. Gonzalez, M. Ostermeier, *J. Mol. Biol.* **2019**, *431*, 1981–1992.

[17] J. D. Bloom, S. T. Labthavikul, C. R. Otey, F. H. Arnold, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 5869–5874.

[18] C. A. Tracewell, F. H. Arnold, *Curr. Opin. Chem. Biol.* **2009**, *13*, 3–9.

[19] B. Csörgo, T. Fehér, E. Tímár, F. R. Blattner, G. Pósfai, *Microb. Cell Fact.* **2012**, *11*, 11.

Manuscript received: March 1, 2023
Accepted manuscript online: April 5, 2023
Version of record online: May 3, 2023