

Microfluidic Screening And Genomic Mutation Identification For Enhancing Cellulase Production in *Pichia Pastoris*

Huiling Yuan

Tianjin Institute of Industrial Biotechnology Chinese Academy of Sciences

Ying Zhou

Wuhan Sunhy Biology Co., Ltd.

Ran Tu

Tianjin Institute of Industrial Biotechnology Chinese Academy of Sciences

Yuping Lin

Tianjin Institute of Industrial Biotechnology Chinese Academy of Sciences

Yufeng Guo

Tianjin Institute of Industrial Biotechnology Chinese Academy of Sciences

Yuanyuan Zhang

Tianjin Institute of Industrial Biotechnology Chinese Academy of Sciences

Qinhong Wang (✉ wang_qh@tib.cas.cn)

Tianjin Institute of Industrial Biotechnology Chinese Academy of Sciences <https://orcid.org/0000-0003-0004-2721>

Research

Keywords: Cellulase, *Pichia pastoris*, High throughput screening (HTS), Whole-cell directed evolution, Droplet-based microfluidics, Whole-genome resequencing

Posted Date: November 23rd, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-1071826/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background

Pichia pastoris is a widely used host organism for heterologous production of industrial enzymes and other proteins, such as cellulase applied in various industries. By developing rational and irrational engineering strategies, great progress has been achieved in improving protein expression in *P. pastoris*. However, due to unknown genomic impact factors, the potential of the *P. pastoris* expression system has not been fully explored. Recently, whole-cell directed evolution, employing iterative rounds of genome-wide diversity generation and high throughput screening (HTS), is considered as a promising strategy in strain improvement at the genome level.

Results

In this study, whole-cell directed evolution of *P. pastoris*, employing atmospheric and room temperature plasma (ARTP) mutagenesis and droplet-based microfluidic high-throughput screening (HTS), was developed to improve heterogenous cellulase production. The droplet-based microfluidic HTS platform based on a cellulase-catalyzed reaction of releasing fluorescence was established to be suitable for methanol-grown *P. pastoris*. The validation experiment showed a positive sorting efficiency of 94.4% at a sorting rate of 300 drops per second. Through five rounds of iterative ARTP mutagenesis and microfluidic screening, cellulase production capacities of mutants in each round were gradually enhanced. Among them, the best mutant strain R5-2 exhibited the cellulase activity of $11,110 \pm 523$ U/L, an about 2-fold increase compared to the starting strain. Whole-genome resequencing analysis further uncovered three accumulated genomic alterations in coding region during iterative ARTP mutagenesis and screening. Intriguingly, the point mutation Rsc1^{V22G} was observed in all the top-performing producers selected from each round, and gene deletion analysis confirmed that Rsc1, a component of the RSC chromatin remodeling complex, might play an important role in cellulase production.

Conclusions

We established a droplet-based microfluidic HTS system, thereby facilitating whole-cell directed evolution of *P. pastoris* for enhancing cellulase production, and meanwhile identified genomic alterations by whole-genome resequencing and analyzed. Our approaches and findings would provide guides to accelerate whole-cell directed evolution of host strains and enzymes of high industrial interest.

Background

Cellulase can catalyze the hydrolysis of cellulose into fermentable monosaccharides and has numerous applications in biofuel production, textile polishing and finishing, paper, laundry, food, and feed processing industries [1–3]. Currently, filamentous fungi genera, such as *Trichoderma*, *Penicillium*, and *Aspergillus*, are the major source of commercial cellulase mixtures [4–6]. However, the end-product inhibition to cellulase expression and activity still remains a challenge [7]. Therefore, to address this

obstacle, heterologous expression of cellulase genes in non-cellulolytic microorganisms becomes a promising strategy [8–10]. Yeast, a simple and effective single-cell eukaryotic organism, is frequently harnessed as a recombinant protein expression system. Among yeast species, the methylotrophic yeast *Pichia pastoris* (syn. *Komagataella phaffii*) has become a widely used host organism for heterologous protein production, and the *P. pastoris* expression technology has been commercially available for many years [11–15]. The success of the *P. pastoris* expression system can be attributed to several advantages: (i) it can be grown to very high cell densities in inexpensive media by relatively easy manipulation of fermentation parameters; (ii) it has a high-level secretion system of heterologous proteins; (iii) it has strong and tightly regulated promoters; and (iv) it can perform higher eukaryotic protein post-translational modifications [16]. To date, more than 5000 proteins have been expressed in *P. pastoris* systems [15]. By applying various rational and irrational engineering strategies, significant progress has been achieved in improving recombinant protein expression by *P. pastoris* [17]. However, it is still hard to further improve recombinant protein production on the basis of high expression levels in industrial strains already in use, which is mostly due to unknown impact factors hidden in its genome [18]. Thus, it is worthy of generating genome-wide mutagenesis and more efficient screening methods, thus discovering more genomic variants associated with heterogenous protein expression in *P. pastoris*.

Whole-cell directed evolution, which employs iterative rounds of diversity generation and screening, is considered as a promising strategy to develop desirable production traits at the genome-wide level [19–21]. Genome-wide diversity generation approaches have been developed to be more efficient and well applied, such as irrational methods including UV mutagenesis, EMS chemical mutagenesis, genome shuffling, atmospheric and room temperature plasma (ARTP) mutagenesis, etc., and rational methods including site-directed RNAi, CRISPR-Cas, recombineering, and transposon insertion, etc. [22–28]. By contrast, the lack of suitable high throughput screening (HTS) methods is still a primary bottleneck [29, 30]. Conventional agar plate- and microtiter plate-based assays are the most widely used screening formats. However, the screening throughput of these formats is limited, and the process is laborious [31]. Fluorescence-activated cell sorting (FACS) can screen individual cells at a rate of up to 10^8 cells per day, but can not be used for secreted metabolites and enzymes. In recent years, droplet-based microfluidics has emerged as a powerful HTS method because of its significantly higher throughput and lower reagent consumption compared to MTP, and its compartmentalization assays in emulsion droplets compared to FACS [32–34]. In this approach, monodisperse picolitre water-in-oil droplets are generated and manipulated (such as droplet splitting, fusion, picoinjection, detection, and sorting) at rates of thousands per second. These droplets can be used as microreactors to compartmentalize individual samples (cell, DNA, protein, chemical/biological reaction, etc.) for reliable quantitative analysis. Through compartmentalization, droplets provide a high-performance platform to enhance the linkage of phenotype and genotype during whole-cell directed evolution [35]. In the past few years, droplet microfluidics has been successfully applied to directed evolution of different microorganisms, including bacteria, yeast, and filamentous fungi, which expressed different enzymes including β -galactosidase, cellulase, α -amylase, aldolase, and esterase [31, 36–41].

In this study, to further enhance heterogenous protein production performances of an industrial recombinant cellulase-producing *P. pastoris* strain, we developed a droplet-based microfluidics HTS system according to a cellulase-catalyzed reaction of releasing fluorescence (Fig. 1a), combined with ARTP mutagenesis method, to conduct whole-cell directed evolution. The droplet-based microfluidic platform including droplet generation, off-chip incubation, reinjection, and sorting, was established to be suitable for methanol-grown *P. pastoris*. After five rounds of iterative ARTP mutagenesis and microfluidic screening, the best cellulase-producing mutant strain R5-2 was obtained, which exhibited a 2.1(±0.1)-fold increase of production and a 1.8(±0.1)-fold increase of yield compared to the starting strain. Furthermore, genome resequencing analysis revealed three accumulated genomic alterations in coding region during iterative mutagenesis and screening. The point mutation Rsc1^{V22G} was observed in all the top-performing producers selected from each round, and gene deletion analysis confirmed that Rsc1 might play an important role in cellulase production. The other two mutations Rvs167^{G9R} and Pmt6^{V57M} were only found in one selected mutant strain and gene deletion analysis suggested the negative but minor effect of Rvs167 and the additive effect of Pmt6 on cellulase production. Taken together, this study developed and tested a high-throughput and efficient strategy of enhancing heterogenous cellulase production in *P. pastoris*, and meanwhile demonstrated some key genomic impact factors affecting heterogenous protein expression, thus providing more guidance to accelerate whole-cell directed evolution of host strains and enzymes of high industrial interest.

Results

Establishment of microfluidic screening for cellulase-producing *P. pastoris*

To monitor cellulase activity in microfluidic droplets, a fluorescence-based method using a fluorogenic substrate fluorescein di-beta-D-cellobioside (FCB) was employed [42]. Cellulase can convert FCB into the compound fluorescein with green fluorescence (Fig. 1a). To evaluate the performance of this FCB-based cellulase activity measurement method, different concentrations of commercial cellulase were incubated with FCB for 48 h in droplets using microtiter plate (MTP) as a control, and subjected to fluorescence detection. As shown in Figure 1b, both droplet-based and microtiter plate (MTP)-based assay formats indicated positive correlations between cellulase activities at the range of 0-4 U/L and fluorescence intensities.

Methanol is usually used as the sole carbon source and inducer for protein expression by recombinant *P. pastoris* strains. Furthermore, the retention of fluorescent marker in water-in-oil droplets is important to the precision of fluorescence-activated droplet sorting. However, the effects of methanol, as an organic solvent, on the stability and fluorescence leakage of the water-in-oil droplet have not yet been evaluated. First, therefore, the droplets containing different concentrations of methanol (0 to 2%, v/v) were formed, and the diameters of droplets were measured by image analysis over time. The results showed that the sizes of droplets have no obvious changes during 48 h incubation (Fig. S1a). Furthermore, microscopy

images showed that the droplets still maintained a high degree of monodispersity (Fig. S1b). Second, to test the effect of methanol on fluorescence leakage, three types of droplets, including nonfluorescent droplets only with BMMY medium, fluorescent droplets with cellulase and FCB, and the binary mixture of non-fluorescent and fluorescent droplets with a 10:1 mixing ratio, were generated, incubated for 48 h, and thereafter subjected to microfluidic measurement of their fluorescence distributions. Compared to the fluorescence distributions of non-fluorescent droplets and fluorescent droplets, histogram of the binary droplet mixture showed a clear separation of non-fluorescent and fluorescent droplets (Fig. 2a-c). During the microfluidic detection, the fluorescence signals of non-fluorescent and fluorescent droplets can be well distinguished (Fig. 2d), which was further confirmed by microscope image of droplets (Fig. 2e). Taken together, methanol, which is required for protein expression by recombinant *P. pastoris* strains, would have no obvious effect on subsequent microfluidic screening.

To determine the optimal incubation time for droplet detection and sorting, the cellulase-producing *P. pastoris* strain SHY169 and FCB were encapsulated in droplets, and their green fluorescence signals were detected by fluorescence microscopy at different time points. After 24 h of cultivation, the signals became detectable, and the signal intensities continued to increase until 40 h (Fig. 3a). Thus, we selected 40 h as the incubation time for droplet sorting.

Lastly, we validated whether the established droplet-based microfluidic platform can be used to sort out *P. pastoris* cells with cellulase production. A sorting experiment was conducted to enrich cellulase-producing SHY169 cells from the artificially mixed library, which was composed of SHY169 and non-cellulase producing GS115 at a 1:99 ratio. The histogram of background fluorescence from droplets at the 0-h timepoint showed only one peak centered on the fluorescence (a.u.) of 0.04 (Fig. 3b). After 40-h incubation, droplets with higher fluorescence were differentiated from the background fluorescence (Fig. 3c), and the 0.1% most fluorescent droplets were sorted out by setting a threshold at the fluorescence (a.u.) of 0.07. Ninety cells from the selected droplets were further validated for cellulase production capacities using flask experiments, and cellulase activity was measured by the 3,5-dinitrosalicylic acid (DNS) reduction method. Meanwhile, ninety cells from the original mixed library were used as a control. Before sorting, one in ninety cells showed cellulase activity, while eighty-five in ninety cells showed cellulase activity after sorting (reds well in Fig. 3d except for negative and positive controls in the first six wells in the first column), resulting in about 94.4% positive sorting efficiency and representing enrichment ratio of 1513 [32]. These results indicated that our droplet-based microfluidic platform could be efficient for the high-throughput sorting of cellulase-producing cells.

Overall, the droplet-based microfluidic platform was established as follows (Fig. 4): (a) encapsulation of single cells together with a fluorogenic substrate FCB in 10 pL droplets (Fig. 4a); (b) off-chip incubation for cell growth and cellulase expression for 40 h at 30°C (Fig. 4b); (c) reinjection of droplets for analysis and sorting based on the fluorescence signal, where about 2×10^6 droplets were sorted at a throughput of 300 droplets per second and 0.01–0.03% droplets with positive signals were collected (Fig. 4c-e); (d) recovery of cells from the sorted droplets.

Microfluidic screening of enhanced cellulase production from ARTP mutagenesis libraries

We next applied the droplet-based microfluidic system for screening cellulase hyperproducers from ARTP-mutagenized *P. pastoris*. To increase the chance of obtaining desirable mutants, five rounds of iterative ARTP mutagenesis and screening were conducted. For each round, the best cellulase producer was selected to be a starting strain for the next round. Microfluidic screening was performed by following the above workflow (Fig. 4). From each round, 200 to 600 droplets with positive green fluorescence were sorted out and subjected to the subsequent primary, secondary and final validations by flask fermentation and 5-L bioreactor fermentation for the best producer. Cellulase production performance of cells was evaluated by two parameters including production and yield. Production is measured in units of cellulase activity per volume, whereas yield is measured in units of cellulase per OD₆₀₀ of biomass. First, about thirty clones from each round were randomly picked from the recovery plates for primary validation (Fig. 5). Validation rate was calculated as the percentage of isolates with higher cellulase production and yield compared to the starting strain. In terms of cellulase production, the first round showed the highest validation rate of 90%, while the fifth round showed the lowest validation rate of 42% (Fig. 5a). In terms of cellulase yield, the first round showed the lowest validation rate of 42%, while the third round showed the highest validation rate of 87% (Fig. 5a). Overall, an average validation rate was around 70%. Furthermore, compared to the starting strain, cellulase production in each of the five rounds were maximumly increased 1.8, 1.4, 1.6, 2.0, and 1.6-fold, respectively, whereas cellulase yield were maximumly increased 1.3, 1.4, 1.7, 1.9, and 2.0-fold (Fig. 5b-f). These results indicated that the iterative ARTP mutagenesis and screening could be an efficient way to improve and enrich cellulase production performance. Additionally, there was no apparent positive relationship between cellulase production and yield for the same isolate. Thus, we selected five isolates from each round, which showed top performance at cellulase production and no decrease in cellulase yield, to be further evaluated (Fig. 5b-f).

Secondary validation results showed that more than one third of mutants seemed to rapidly degenerate (Fig. S2), which consisted with the previously reported genome instability of mutagenized *P. pastoris* [15, 43]. Notably, some mutants from the third and fourth rounds, whose primary validation rates were relatively high (Fig. 5a), such as R3-4, R4-5, and R4-12, maintained the significance of improved performance (Fig. S2). Thus, we selected one best producers from the first and second rounds, top three producers from the third and fourth rounds as well as top two producers for final validation. Eventually, compare to the wild type stain SHY169, most of the selected mutants except for R4-6 exhibited significantly increased cellulase production by 1.5 to 2.1-fold as well as yield by 1.3 to 1.8-fold (Fig. 6a). Among them, the mutant stain R5-2, which was obtained from the five round, showed the best cellulase production performance with a 2.1(±0.1)-fold increase of production and a 1.8(±0.1)-fold increase of yield (Fig. 6a). High-density fermentation experiments in a 5-L bioreactor were conducted to further compare cellulase production capacities between the mutant stain R5-2 and the wild type strain SHY169. The fermentation profiles showed that the mutant R5-2 had a faster and higher cell growth than the wild type SHY169 (Fig. 6b), where the final cell growth of R5-2 increased 1.3-fold than that of SHY169.

Correspondingly, R5-2 exhibited faster and higher cellulase production than that of SHY169, where R5-2 produced a 1.9-fold increase of cellulase ($11,110 \pm 523$ U/L) by a 1.7-fold increase of rate (74.6 U/L/h) compared to SHY169 (Fig. 6b).

Elucidation of genomic variants in mutant strains with enhanced cellulase production

To uncover genomic alterations accumulated during iterative ARTP mutagenesis, the ten selected mutant strains according to the final validation results (Fig. 6a) and the wild-type strain were subjected to whole-genome resequencing. By comparing with the reference genome of strain GS115, different SNPs and InDels were detected in the mutant strains versus the starting strain in each round (Additional file 2: Dataset S1). Here, only mutations in coding regions were considered. Total three SNPs (positions: 2290916 and 214209 on Chromosome 2, 1516266 on Chromosome 4) were found in coding regions, thereby causing nonsynonymous amino acid changes in three proteins, including Rsc1 (V22G, valine to glycine at position 22), Rvs167 (G9R, glycine to arginine at position 9) and Pmt6 (V57M, valine to methionine at position 57). Rsc1 (GeneBank accession No. XP_002492164) is a component of the RSC chromatin remodeling complex [44], and its homolog in *Saccharomyces cerevisiae* was reported to be involved in general transcription regulation [45]. Rvs167 (GeneBank accession No. XP_002491000) is an actin-associated protein, and its homolog in *S. cerevisiae* was reported to play roles in endocytosis and exocytosis as well as interact with proteins involved in ER to Golgi vesicle trafficking [46, 47]. Pmt6 is one of protein-O-mannosyltransferases, which was previously characterized to play a minor role in O-glycosylation in *P. pastoris* [48].

To further trace their accumulations during iterative mutagenesis, these three mutations were mapped onto the mutagenesis lineage (Fig. 7a). The mutation of Rsc1^{V22G} was observed to be generated in the first round of ARTP mutagenesis and passed on to all the next generations. The stable inheritance of Rsc1^{V22G} in high cellulase-producing mutants implicated that Rsc1 might play an important role in cellulase production. Additionally, besides Rsc1^{V22G}, Rvs167^{G9R} was found in the mutant R3-24 selected from the third round, which showed the lowest cellulase production among the three selected strains (Fig. 6a). Thus, it seemed that Rvs167 might have a minor or even negative effect on cellulase production. Interestingly, besides Rsc1^{V22G}, the best cellulase producer R5-2 selected from the fifth round was found to accumulate the other mutation Pmt6^{V57M}, suggesting its additive beneficial effect on cellulase production.

Next, gene deletion method was employed to further validate the effects of Rsc1, Rvs167, and Pmt6 on cellulase production (Fig. 7b). All the selected mutant strains were found to have the point mutation in *RSC1* by genome sequencing (Fig. 7a). Deletion of *RSC1*, either in the wild-type strain SHY169 or some mutant strains, such as R1-3 from the first round, R3-24 from the third round, and R5-2 from the fifth round, resulted in significantly decreased cellulase production. This result suggested that *RSC1* might play an important role in cellulase production. Deletion of *RVS167* resulted in significantly increased cellulase production in the wild-type strain SHY169, but not in the mutant strain R3-24 harboring the

mutant Rvs167^{G9R}. This result implicated that *RVS167* might play a negative role in cellulase production, but relatively minor. *PMT6* point mutation was found in the mutant strain R5-2. Deletion of *PMT6* resulted in a significantly 0.85-fold decrease in cellulase production by R5-2, but its impact was slightly lower than deletion of *RSC1*, which showed a 0.71-fold decrease. By contrast, the effects of *PMT6* and *RSC1* on cellulase production in the wild-type strain SHY169 were similar, where 0.88-fold and 0.82-fold decreases were observed, respectively. These results suggested that *PMT6* might play a positive but minor role in cellulase production. Overall, gene deletion experiments further confirmed the potential functions of Rsc1, Rvs167, and Pmt6, which accumulated point mutations during iterative ARTP mutagenesis, in cellulase production.

Discussion

P. pastoris is one of the most attractive hosts for the heterogenous and large-scale production of cellulases. Due to limited understanding of complicated cellular metabolism, whole-cell directed evolution has played a significant role in strain improvement to increase enzyme production. However, the inefficiency of conventional plate-based screening is a major bottleneck. In the past decade, droplet microfluidics has emerged as a powerful tool for high throughput screening. A previous study reported that a gel microdroplet-based high-throughput screening method was employed to screen *P. pastoris* strains with improved xylanase production at a throughput of up to 10⁸ variants per day [49]. Compared to gel microdroplet, the generation and manipulation (including splitting, fusion, trapping, injection of reagent, and sorting) of water-in-oil droplet were more flexible [50]. Here, we present a water-in-oil droplet microfluidic method for high-throughput screening of mutant *P. pastoris* strain libraries for improved cellulase production.

Compared to flow cytometry-based high-throughput screening, an important advantage of droplet-based screening method is that each droplet serving as an individual microreactor to link genotype (the enzyme-encoding gene) and phenotype (the reaction product) through compartmentalization. This linkage relies on the integrity of the droplet. Therefore, the droplet stability and fluorescein leakage were tested in the presence of methanol in the aqueous phase, and the results showed that after 48 h off-chip incubation, the droplets remained stable and monodisperse in the methanol concentration range of 0.5%-2%, and the fluorescein leakage was insignificant in the presence of 1% methanol in droplets (Fig. S1). Other factors that impact droplet integrity have previously been reported and should be noted, including the choice of oil-surfactant pairs, droplet incubation conditions, and the addition of Bovine Serum Albumin prevents leaking, etc.

The droplet microfluidic screening platform provides higher throughputs and lower reagent consumption compared to traditional microplate-based screening. In this study, *P. pastoris* single cells were encapsulated in 10 pL droplets at a rate of 3000 droplets per second, and droplets were sorted at 300 droplets per second based on their fluorescence intensity (Fig. 4). The total reagent consumption was less than 2 mL in each round of screening. After five sequential rounds of mutagenesis and selection, several mutants with improved cellulase production were successfully isolated and the best-performing

clone exhibited a more than 2-fold increase in cellulase production (Fig. 6). Furthermore, the microfluidic screening system could potentially be adapted for engineering and screening strains for the expression of other enzymes or even other metabolites as long as they can be linked to a fluorescent signal.

In addition, with the rapid development of high-throughput sequencing technology and decrease in sequencing costs, whole-genome sequencing technology has now enabled us to analyze genotype-phenotype correlations of mutants with improved performance, identify the genetic changes and mechanisms for the altered phenotype, and combine with accurate genetic modification techniques, to rationally guide and optimize strain improvement. Huang et al. reported used droplet microfluidics screening of whole-genome yeast mutant libraries for improved α -amylase secretion, and performed whole-genome sequencing of several isolated strains, and identified 330 mutations that may contribute to improved protein secretion [31]. Balasubramanian et al. reported used random mutagenesis and droplet-based microfluidic to successfully improve β -glucosidase secretion in *C. glutamicum*, and whole-genome resequencing and functional enrichment analysis found mutations involved in protein synthesis and secretion relevant biological processes [51]. Similar to previous studies, we also sequenced the whole-genomes of the improved cellulase production mutants from each library together with the original strain, several mutations that may contribute to the improved α -amylase secretion were identified (Fig. 7). These mutations identified in this study can potentially be transferred to other platform strains by reverse metabolic engineering for the production of other proteins. The combined use of microfluidics screening and whole-genome sequencing could broadly facilitate the design of novel cell factories.

Conclusions

In summary, we developed a droplet-based microfluidics platform for high throughput screening of *P. pastoris* strains with improved cellulase production at a throughput of up to 300 drops per second. We performed five rounds of whole-genome mutagenesis and selection, obtained a *P. pastoris* mutant strain with about 2-fold increase in cellulase production. Mutated genes identified by whole-genome sequencing were demonstrated that affecting heterogenous protein expression. Furthermore, the droplet-based microfluidic HTS platform might be also applicable to accelerate the whole-cell directed evolution of host strains of other secreted enzymes.

Methods

Strains, media, and culture conditions

The strains used in this study are listed in Additional file 1: Table S1. *P. pastoris* SHY169 with integrated cellulase gene and derived from GS115 was provided by Wuhan Sunhy Biology Co., Ltd (Wuhan, China) and used as the starting strain for ARTP mutagenesis. To validate the effects of mutant genes found in ARTP mutant strains on cellulase production, gene deletion strains were generated by using PCR-mediated homologous recombination method [52]. The PCR primers used in this study are listed in Additional file 1: Table S2. Hygromycin-resistance gene was PCR amplified from the plasmid pRS426-

hphB, and its expression was controlled under *S. cerevisiae* *TEF1* promoter and terminator [53]. The PCR fragments containing hygromycin-resistance expression cassettes flanked by 1000-bp homologous sequences to the target genes were obtained using fusion PCR. The 1000-bp homologous sequences upstream and downstream of the target genes were PCR amplified from SHY169 genomic DNA. Yeast transformation was performed by using the electrotransformation method [54]. Positive transformants were selected on hygromycin selective plates, which were YPD agar plates containing 200 µg/ml hygromycin B. Successful gene deletions were verified by diagnostic PCR reactions with primers designed on the target genes as well as hygromycin-resistance cassette (for primers, see Additional file 1: Table S2).

The YPD medium contained, per liter, 10 g yeast extract, 20 g peptone, and 20 g glucose. The BMGY liquid medium contained, per liter, 10 g yeast extract, 20 g peptone, 2.3 g K₂HPO₄, 11.8g KH₂PO₄, pH 6.0, 13.4 g yeast nitrogen base (YNB), 0.4 mg biotin, and 10 ml glycerol. The BMMY medium contained as per BMGY except with the indicated amount of methanol instead of the glycerol. The basal fermentation medium for 5-L jar fermenter included, per liter, 27 g NH₄H₂PO₄, 3.7 g KH₂PO₄, 13.2 g MgSO₄, 13.5 g K₂SO₄, 0.73 g CaSO₄, 1.06 g KOH, 0.44 g Biotin, and 40 g Glycerol. Unless stated otherwise, cells were cultured in YPD or BMGY media at 30 °C, 250 rpm. Cells for inducing cellulase production were grown in BMMY medium at 28 °C, 250 rpm. Fed-fermentation was conducted as follows. Cells were cultured in 60 ml YPD at 28 °C, 150 rpm in 500 ml baffled shaker flasks for 48 h, then transferred to the 5-L jar fermenter containing 2 L of the basal fermentation medium. Methanol was supplemented according to the fermentation process to ensure that methanol concentration was maintained at 0.5 to 1% (v/v).

Fluorescence-based cellulase activity measurement method

To establish a method of monitoring the cellulase activity in droplets, a fluorogenic substrate fluorescein di-beta-D-cellobioside (FCB, AAT bioquest), was used [55]. Under cellulase treatment, FCB releases the green-fluorescent compound fluorescein, whose intensity can be titrated with fluorescence detection (Ex 473 nm/Em 520 nm). To verify the detection range of this fluorescence-based cellulase activity measurement method, activity of commercial cellulase (a commercial product produced by SHY169 strain, Wuhan SunHY Biological Co., Ltd) was assayed in the microfluidic droplets, and the microtiter plate (MTP)-based assay format was used as control. Reactions contained 20 µM FCB and the indicated amounts of cellulase (1-4 U/L), which were dissolved in BMMY medium with 1% methanol to mimic the actual droplet screening. The reaction mixture was rapidly encapsulated into the droplets or pipetted into microplate wells, respectively. After incubating at 30 °C for 48 hours, the fluorescence was measured using the custom droplet screening device in our lab for the droplet-based assay format and using the SpectraMax M2e Microplate Reader (Molecular Devices Corporation) for the MTP-based assay format, respectively.

Droplet-based microfluidics

Fabrication of microfluidic chips

Microfluidic chips were fabricated in poly(dimethylsiloxane) and glass by using standard soft lithographic methods [35]. The chips were used for droplet generation and sorting, respectively. The chips had channel heights about 25 μm for droplet-making or sorting, and 0.5-mm diameter holes for channel inlets and outlets. For the sorting chip, microelectrodes were fabricated by heating the chip to 95 °C and pushing low-melting-temperature solder wire into the electrode channels. A hydrophobic surface treatment was applied to the microfluidic channel walls by injecting Aquapel and then flushing with pressurized air.

Optical setup, data acquisition and control system

The optical setup for droplet fluorescence detection consisted of a 100 mW, 473 nm high stability blue laser and a Nikon TI-U inverted microscope. The laser beam was passed through a beam expander, then guided into the backport of the microscope. The laser light inside the microscope was reflected by a dichroic beam splitter (DBS) and then focused on the sorting chip through the objective lens. The fluorescent emission from each droplet was passed back along the path of the laser and separated by the DBS to the photomultiplier tube (PMT, Hamamatsu Photonics), which was mounted to the microscope's side port. A high-speed video camera (Fastec HiSpec 1, Fastec Imaging, San Diego, CA, USA) was mounted on the microscope for the imaging of droplet formation, reinjection, and sorting. Data acquisition (DAQ) and control were conducted with a DAQ card (National Instruments) executing a program written in LabVIEW software (National Instruments). When the droplet fluorescence intensity exceeds a defined threshold, the DAQ card provided a signal to a high voltage amplifier (TREK 609E-6), which is connected to the electrodes of the sorting chip, to sort the corresponding droplet.

Droplet generation and microfluidic screening

Monodispersed droplets were generated using a droplet-making chip by flow-focusing the aqueous phase with two streams of fluorinated oil (HFE-7500, 3M) containing 2% (wt/wt) surfactant (Raindance Technologies) [56]. The aqueous and oil phases were loaded into 1-ml disposable syringes (BD Biosciences), respectively, and pumped into the droplet-making chip using the syringe pumps (Harvard Apparatus Inc.). Syringes were connected to the droplet-making chip using PTFE tubing with an internal diameter of 0.38 mm and an external diameter of 0.76 mm. The syringe pumps were operated with the aqueous flow rates of 100 $\mu\text{L}/\text{h}$ and the oil flow rates of 300 $\mu\text{L}/\text{h}$, to produce 10-pL droplets at a rate of 3000 droplets per second. The resulting droplets were collected into a 1-mL syringe and incubated for the desired time at 30 °C. After incubation, the droplets were reinjected into a microfluidic sorting chip at a flow rate of 10 $\mu\text{L}/\text{h}$ and spaced by HFE-7500 oil without surfactant infused at 200 $\mu\text{L}/\text{h}$. The emitted fluorescence signal of each droplet was detected by the optical setup. When the droplet fluorescence intensity exceeded a defined threshold, the electrodes were activated, and the droplets were subsequently sorted to the sorting channel, whereas unwanted droplets flowed into the waste channel. The sorted droplets were collected into an Eppendorf tube for further experiments.

Assessment of droplet stability and fluorescence leakage

To test the effects of methanol on the droplet stability and fluorescence leakage, the BMMY medium containing methanol was tested as the aqueous phase for generating droplets. For testing the droplet stability, the droplets containing different concentrations of methanol (0, 0.5%, 1%, or 2%; v/v) were collected into the syringes and incubated at 30 °C for 48 h. Droplets were imaged using Lecia DM5000B microscope (Lecia Microsystems) and compared at the time of droplet formation and during the incubation. The droplet sizes were measured by ImageJ analysis as previously described [57]. To determine fluorescence leakage, three types of droplets, including fluorescent and non-fluorescent droplets as well as their binary mixture, were generated and collected into syringes, respectively. Fluorescent droplets were generated from a 40 h reaction solution of cellulase in BMMY medium with 1% methanol and 20 µM fluorogenic substrate FCB. Non-fluorescent droplets only contained BMMY medium with 1% methanol. The binary mixture contained 10 µL of fluorescent droplets and 100 µL of non-fluorescent droplets. All the droplets in syringes were incubated at 30 °C for 48 h, and separately reinjected into the sorting chip. And then their fluorescence distributions were counted.

Optimal droplet incubation time and verification of the droplet-based microfluidic platform

Optimal droplet incubation time was determined by growing the cellulase-producing strain SHY169 in droplets, which were detected for green fluorescence signals by imaging with Lecia DM5000B fluorescence microscope (Lecia Microsystems) every 8 hours for 48 hours. The droplet-based microfluidic platform was verified by measuring the sorting efficiency of an artificial cell library, which consisted of cellulase-producing and non-cellulase-producing strains. Cellulase-producing strain SHY169 and non-cellulase-producing strain GS115 were grown in BMMY medium overnight at 30 °C, 250 rpm. Cells were pelleted by centrifuging at 3000× g for 2 min, washed twice and resuspended in BMMY medium, and sonicated for 3 × 10 seconds at 40 W to avoid cell adhesion. Cell suspensions of SHY169 and GS115 were mixed at a 1:99 ratio, and then diluted to a concentration of 2×10^7 cells per ml in BMMY medium containing 20 µM FCB and 10% (v/v) density-matching solution (OptiPrep). The mixture was encapsulated in droplets at a cell to droplet ratio of 0.2 according to the law of Poisson distribution, and incubated off-chip at 28 °C for 40 hours. The droplets were reinjected into the sorting chip, and sorted according to their fluorescence intensities at a rate of 300 droplets per second. Cells were recovered from the sorted droplets by adding 50 µL of 1H,1H,2H,2H-perfluorooctanol and gently vortexing to break the droplets, and spread on YPD agar plate to allow growth for 2 days. Ninety colonies were randomly picked and used for cellulase activity measurement using the 3,5-dinitrosalicylic acid (DNS) reduction method [5].

ARTP mutagenesis and microfluidic screening

ARTP was employed to iteratively mutagenize cellulase-producing strain SHY169. A single colony from a fresh plate was inoculated into YPD medium and grown overnight at 30°C, 250 rpm. Cells were pelleted and washed twice with sterile saline, and diluted to $OD_{600}=1$ (corresponded to 1×10^8 cells per ml). Subsequently, 10 μ L cell suspension was spread on a sterilized stainless-steel plate and subjected to ARTP treatment for 30 s. The operating parameters were as follows, including a radio frequency power input of 100 W, a helium gas flow rate of 10 SLM (standard liters per minute), and a 2 mm distance between the plasma torch nozzle exit and the sample plate. The ARTP mutagenesis procedure was repeated five times. All cells from the five ARTP treated plates were collected in a new sterile tube containing 450 μ L of YPD medium, and spread on YPD agar plates. The agar plates were incubated at 30 °C until colonies formed.

Cells were scraped from the agar plates, washed, and resuspended in fresh BMMY medium. Cell suspension was sonicated for 3×10 s at 40 W and filtered using an 8- μ m filter (Whatman) to avoid cell adhesion. OD_{600} was measured, and cell suspension was diluted to 2×10^7 cells per ml to achieve a cell-to-droplet ratio at 0.2. The diluted cell suspension was subsequently encapsulated into droplets together with 20 μ M fluorogenic substrate FCB. The generated droplets were collected into a 1-ml syringe preloaded with 100 μ L HFE-7500 oil, and then incubated at 30 °C for 40 h. Next, the droplets were sorted based on the resulting fluorescence using the sorting chip. The sorted droplets were collected in a sterile Eppendorf tube with 200 μ L fresh YPD medium. Cells were recovered from the sorted droplets, spread on YPD agar plates, and incubated at 30 °C until colonies formed. About thirty colonies were randomly picked and used for further validating cellulase production by primary validation of flask fermentation, and then the isolates with the top five high performance were applied to secondary validation. The mutant strain with the stable and highest cellulase production was subjected to the next round of ARTP mutagenesis and microfluidic screening.

Characterisation of cellulase production by selected mutants

Cellulase production by selected mutants was characterized by flask fermentation and 5-L bioreactor as specified in the text. The selected colonies from every round of ARTP mutagenesis and microfluidic screening were inoculated in 2 ml BMGY medium, and grown to the logarithmic period at 30°C, 250 rpm. Cells were harvested by centrifuging at 3000 \times g for 5 minutes at room temperature and resuspended to an OD_{600} of 1.0 in 10 ml BMMY medium containing 1% methanol in a 50-mL flask. After incubation for 40 h at 28°C, 250 rpm to induce cellulase production, OD_{600} of each fermentation broth was measured, and cellulase activity of fermentation supernatant was analyzed using the DNS reduction method. Finally, the top performer from flask fermentation was further analyzed for cellulase production in 5-L batch fermentation using the starting strain SHY169 as a control.

Genome resequencing and data analysis

The wild-type strain SHY169 and ten selected mutant strains with enhanced cellulase production were grown in 200 ml liquid YPD media at 30°C to exponential phase. Cells were harvested by centrifuging for

5 minutes at 4000 x g. The supernatants were discarded, and the pellets were stored at -80°C until further use for genome resequencing. Genomic DNA isolation and the sequencing libraries were constructed and sequenced on Illumina HiSeq Xten-PE150 using 150-bp paired-end sequencing by Novogene (Beijing, China). A mean of 9.2 million 150-bp clean reads was generated for each library. The genome sequencing raw data were deposited in the NCBI Sequence Read Archive under the accession number PRJNA772571. Sequence analysis and variant detection were performed as previously described [58]. The *P. pastoris* GS115 genome as a reference was downloaded from RefSeq at the NCBI (sequence assembly version ASM174695v1, RefSeq assembly accession: GCA_001746955.1). The Genome Analysis Toolkit (GATK v3.5) Best Practices pipeline was employed to detect single nucleotide polymorphisms (SNPs) and insertion/deletion (InDels) [59, 60]. The cleaned reads were mapped to the reference genome using the mapping tools BWA-mem (version 0.7.13) [61]. Variant callings were performed using GATK HaplotypeCaller. Variant annotation was conducted using the package ANNOVAR [62]. The genomic variants including SNPs and InDels as well as the protein-altering variants in comparisons between the starting strains and the subsequent selected strain during the iterative ARTP mutagenesis and microfluidic screening are available in supplementary material (Additional file 2: Dataset S1; Additional file 3: Dataset S2).

Declarations

Acknowledgements

Not applicable.

Authors' contributions

Qinhong Wang conceived and supervised this project. Huiling Yuan, Ying Zhou and Ran Tu performed the experiments. Huiling Yuan, Yuping Lin, Yufeng Guo and Yuanyuan Zhang analysed the data. Huiling Yuan and Qinhong Wang wrote the manuscript. All authors read and approved the final manuscript.

Funding

This work was supported by Tianjin Synthetic Biotechnology Innovation Capacity Improvement Project (TSBICIP-KJGG-004), the Instrument Developing Project of the Chinese Academy of Sciences (YJKYYQ20170023) and Industrial Synthetic Biology Innovation Team support funding for Qinhong Wang and Yuping Lin.

Availability of data and materials

All data generated or analyzed during this study are included in this published article and its additional files.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

Ying Zhou is an employee of Wuhan Sunhy Biology Co., Ltd. Materials disclosed in this paper are also subject to one or more patent filings.

References

1. Kuhad RC, Gupta R, Singh A. Microbial cellulases and their industrial applications. *Enzyme Res.* 2011,2011:280696.
2. Garvey M, Klose H, Fischer R, Lambertz C, Commandeur U. Cellulases for biomass degradation: comparing recombinant cellulase expression platforms. *Trends Biotechnol.* 2013,31(10):581–93.
3. Ejaz U, Sohail M, Ghanemi A. Cellulases: From bioactivity to a variety of industrial applications. *Biomimetics (Basel).* 2021,6(3):44.
4. Gusakov AV. Alternatives to *Trichoderma reesei* in biofuel production. *Trends Biotechnol.* 2011,29(9):419–25.
5. Gao F, Hao Z, Sun X, Qin L, Zhao T, Liu W, Luo H, Yao B, Su X. A versatile system for fast screening and isolation of *Trichoderma reesei* cellulase hyperproducers based on DsRed and fluorescence-assisted cell sorting. *Biotechnol Biofuels.* 2018,11:261.
6. Bischof RH, Ramoni J, Seiboth B. Cellulases and beyond: the first 70 years of the enzyme producer *Trichoderma reesei*. *Microb Cell Fact.* 2016,15(1):106.
7. Percival Zhang YH, Himmel ME, Mielenz JR. Outlook for cellulase improvement: screening and selection strategies. *Biotechnol Adv.* 2006,24(5):452–81.
8. Mazzoli R, Lamberti C, Pessione E. Engineering new metabolic capabilities in bacteria: lessons from recombinant cellulolytic strategies. *Trends Biotechnol.* 2012,30(2):111–9.
9. Hasunuma T, Kondo A. Development of yeast cell factories for consolidated bioprocessing of lignocellulose to bioethanol through cell surface engineering. *Biotechnol Adv.* 2012,30(6):1207–18.
10. Lambertz C, Garvey M, Klinger J, Heesel D, Klose H, Fischer R, Commandeur U. Challenges and advances in the heterologous expression of cellulolytic enzymes: a review. *Biotechnol Biofuels.* 2014,7(1):135.
11. Karbalaee M, Rezaee SA, Farsiani H. *Pichia pastoris*: A highly successful expression system for optimal synthesis of heterologous proteins. *J Cell Physiol.* 2020,235(9):5867–81.
12. Yang Z, Zhang Z. Engineering strategies for enhanced production of protein and bio-products in *Pichia pastoris*: A review. *Biotechnol Adv.* 2018,36(1):182–95.

13. Juturu V, Wu JC. Heterologous protein expression in *Pichia pastoris*: latest research progress and applications. *Chembiochem*. 2018,19(1):7–21.
14. Macauley-Patrick S, Fazenda ML, McNeil B, Harvey LM. Heterologous protein production using the *Pichia pastoris* expression system. *Yeast*. 2005,22(4):249–70.
15. Schwarzhans JP, Luttermann T, Geier M, Kalinowski J, Friehs K. Towards systems metabolic engineering in *Pichia pastoris*. *Biotechnol Adv*. 2017,35(6):681–710.
16. Cregg JM, Cereghino JL, Shi J, Higgins DR. Recombinant protein expression in *Pichia pastoris*. *Mol Biotechnol*. 2000,16(1):23–52.
17. Peña DA, Gasser B, Zanghellini J, Steiger MG, Mattanovich D. Metabolic engineering of *Pichia pastoris*. *Metab Eng*. 2018,50:2–15.
18. Dikicioglu D, Wood V, Rutherford KM, McDowall MD, Oliver S. Improving functional annotation for industrial microbes: a case study with *Pichia pastoris*. *Trends Biotechnol*. 2014,32(8):396–9.
19. Si T, Lian J, Zhao H. Strain development by whole-Cell directed evolution. In: Alcalde M, Editor. *Directed enzyme evolution: advances and applications*. Springer International Publishing; 2017. p.173–200.
20. Winkler JD, Kao KC. Recent advances in the evolutionary engineering of industrial biocatalysts. *Genomics*. 2014,104(6 Pt A):406–11.
21. Markel U, Essani KD, Besirlioglu V, Schiffels J, Streit WR, Schwaneberg U. Advances in ultrahigh-throughput screening for directed enzyme evolution. *Chem Soc Rev*. 2020,49(1):233–62.
22. Si T, Luo Y, Bao Z, Zhao H. RNAi-assisted genome evolution in *Saccharomyces cerevisiae* for complex phenotype engineering. *ACS Synth Biol*. 2015,4(3):283–91.
23. Stovicek V, Borodina I, Forster J. CRISPR–Cas system enables fast and simple genome editing of industrial *Saccharomyces cerevisiae* strains. *Meta Eng Commun*. 2015,2:13–22.
24. Ellis HM, Yu D, DiTizio T, Court DL. High efficiency mutagenesis, repair, and engineering of chromosomal DNA using single-stranded oligonucleotides. *Proc Natl Acad Sci U S A*. 2001,98(12):6742–6.
25. Wang HH, Isaacs FJ, Carr PA, Sun ZZ, Xu G, Forest CR, Church GM. Programming cells by multiplex genome engineering and accelerated evolution. *Nature*. 2009,460(7257):894–898.
26. Sauer U. Evolutionary engineering of industrially important microbial phenotypes. *Adv Biochem Eng Biotechnol*. 2001,73:129–69.
27. Zhang X, Zhang XF, Li HP, Wang LY, Zhang C, Xing XH, Bao CY. Atmospheric and room temperature plasma (ARTP) as a new powerful mutagenesis tool. *Appl Microbiol Biotechnol*. 2014,98(12):5387–96.
28. Zhang X, Zhang X, Xu G, Zhang X, Shi J, Xu Z. Integration of ARTP mutagenesis with biosensor-mediated high-throughput screening to improve L-serine yield in *Corynebacterium glutamicum*. *Appl Microbiol Biotechnol*. 2018,102(14):5939–51.

29. Zeng W, Guo L, Xu S, Chen J, Zhou J. High-throughput screening technology in industrial biotechnology. *Trends Biotechnol.* 2020,38(8):888–906.
30. Sarnaik A, Liu A, Nielsen D, Varman AM. High-throughput screening for efficient microbial biotechnology. *Curr Opin Biotechnol.* 2020,64:141–50.
31. Huang M, Bai Y, Sjostrom SL, Hallstrom BM, Liu Z, Petranovic D, Uhlen M, Joensson HN, Andersson-Svahn H, Nielsen J. Microfluidic screening and whole-genome sequencing identifies mutations associated with improved protein secretion by yeast. *Proc Natl Acad Sci U S A.* 2015,112(34):E4689-96.
32. Baret J-C, Miller OJ, Taly V, Ryckelynck M, El-Harrak A, Frenz L, Rick C, Samuels ML, Hutchison JB, Agresti JJ et al. Fluorescence-activated droplet sorting (FADS): efficient microfluidic cell sorting based on enzymatic activity. *Lab Chip.* 2009,9(13):1850–8.
33. Agresti JJ, Antipov E, Abate AR, Ahn K, Rowat AC, Baret JC, Marquez M, Klivanov AM, Griffiths AD, Weitz DA. Ultrahigh-throughput screening in drop-based microfluidics for directed evolution. *Proc Natl Acad Sci U S A.* 2010,107(9):4004–9.
34. Yang J, Tu R, Yuan H, Wang Q, Zhu L. Recent advances in droplet microfluidics for enzyme and cell factory engineering. *Crit Rev Biotechnol.* 2021:1–23.
35. Mazutis L, Gilbert J, Ung WL, Weitz DA, Griffiths AD, Heyman JA. Single-cell analysis and sorting using droplet-based microfluidics. *Nat Protoc.* 2013,8(5):870–91.
36. Sjostrom SL, Bai Y, Huang M, Liu Z, Nielsen J, Joensson HN, Andersson Svahn H. High-throughput screening for industrial enzyme production hosts by droplet microfluidics. *Lab Chip.* 2014,14(4):806–13.
37. Colin PY, Kintsjes B, Gielen F, Miton CM, Fischer G, Mohamed MF, Hyvonen M, Morgavi DP, Janssen DB, Hollfelder F. Ultrahigh-throughput discovery of promiscuous enzymes by picodroplet functional metagenomics. *Nat Commun.* 2015,6:10008.
38. Qiao Y, Zhao X, Zhu J, Tu R, Dong L, Wang L, Dong Z, Wang Q, Du W. Fluorescence-activated droplet sorting of lipolytic microorganisms using a compact optical system. *Lab Chip.* 2017,18(1):190–6.
39. Ma F, Chung MT, Yao Y, Nidetz R, Lee LM, Liu AP, Feng Y, Kurabayashi K, Yang GY. Efficient molecular evolution to generate enantioselective enzymes using a dual-channel microfluidic droplet screening platform. *Nat Commun.* 2018,9(1):1030.
40. He R, Ding R, Heyman JA, Zhang D, Tu R. Ultra-high-throughput picoliter-droplet microfluidics screening of the industrial cellulase-producing filamentous fungus *Trichoderma reesei*. *J Ind Microbiol Biotechnol.* 2019,46(11):1603–10.
41. Tu R, Zhang Y, Hua E, Bai L, Huang H, Yun K, Wang M. Droplet-based microfluidic platform for high-throughput screening of *Streptomyces*. *Commun Biol.* 2021,4(1):647.
42. Körfer G, Pitzler C, Vojcic L, Martinez R, Schwaneberg U. In vitro flow cytometry-based screening platform for cellulase engineering. *Sci Rep.* 2016,6(1):26128.
43. Zhu T, Guo M, Sun C, Qian J, Zhuang Y, Chu J, Zhang S. A systematical investigation on the genetic stability of multi-copy *Pichia pastoris* strains. *Biotechnol Lett.* 2009,31(5):679–84.

44. De Schutter K, Lin YC, Tiels P, Van Hecke A, Glinka S, Weber-Lehmann J, Rouze P, Van de Peer Y, Callewaert N. Genome sequence of the recombinant protein production host *Pichia pastoris*. *Nat Biotechnol*. 2009,27(6):561–6.
45. Yu F, Imamura Y, Ueno M, Suzuki SW, Ohsumi Y, Yukawa M, Tsuchiya E. The yeast chromatin remodeler Rsc1-RSC complex is required for transcriptional activation of autophagy-related genes and inhibition of the TORC1 pathway in response to nitrogen starvation. *Biochem Biophys Res Commun*. 2015,464(4):1248–53.
46. Friesen H, Colwill K, Robertson K, Schub O, Andrews B. Interaction of the *Saccharomyces cerevisiae* cortical actin patch protein Rvs167p with proteins involved in ER to Golgi vesicle trafficking. *Genetics*. 2005,170(2):555–68.
47. Youn JY, Friesen H, Kishimoto T, Henne WM, Kurat CF, Ye W, Ceccarelli DF, Sicheri F, Kohlwein SD, McMahon HT et al. Dissecting BAR domain function in the yeast Amphiphysins Rvs161 and Rvs167 during endocytosis. *Mol Biol Cell*. 2010,21(17):3054–69.
48. Nett JH, Cook WJ, Chen MT, Davidson RC, Bobrowicz P, Kett W, Brevnova E, Potgieter TI, Mellon MT, Prinz B et al. Characterization of the *Pichia pastoris* protein-O-mannosyltransferase gene family. *PLoS One*. 2013,8(7):e68325.
49. Ma C, Tan ZL, Lin Y, Han S, Xing X, Zhang C. Gel microdroplet-based high-throughput screening for directed evolution of xylanase-producing *Pichia pastoris*. *J Biosci Bioeng*. 2019,128(6):662–8.
50. Kintses B, Vliet L, Devenish S, Hollfelder F. Microfluidic droplets: new integrated workflows for biological experiments. *Current Opinion Chem Biol*. 2010,14(5):548–55.
51. Balasubramanian S, Chen J, Wigneswaran V, Bang-Berthelsen CH, Jensen PR. Droplet-based microfluidic high throughput screening of *Corynebacterium glutamicum* for efficient heterologous protein production and secretion. *Front Bioeng Biotechnol*. 2021,9:668513.
52. Winzeler EA, Shoemaker DD, Astromoff A, Liang H, Anderson K, Andre B, Bangham R, Benito R, Boeke JD, Bussey H et al. Functional characterization of the *S. cerevisiae* genome by gene deletion and parallel analysis. *Science*. 1999,285(5429):901–6.
53. Wang Z, Qi Q, Lin Y, Guo Y, Liu Y, Wang Q. Correction to: QTL analysis reveals genomic variants linked to high-temperature fermentation performance in the industrial yeast. *Biotechnol Biofuels*. 2019,12(1):83.
54. Lin-Cereghino J, Wong WW, Xiong S, Giang W, Luong LT, Vu J, Johnson SD, Lin-Cereghino GP. Condensed protocol for competent cell preparation and transformation of the methylotrophic yeast *Pichia pastoris*. *Biotechniques*. 2005,38(1):44, 46, 48.
55. Körfer G, Pitzler C, Vojcic L, Martinez R, Schwaneberg U. In vitro flow cytometry-based screening platform for cellulase engineering. *Sci Rep*. 2016,6:26128.
56. Holtze C, Rowat AC, Agresti JJ, Hutchison JB, Angile FE, Schmitz CH, Koster S, Duan H, Humphry KJ, Scanga RA et al. Biocompatible surfactants for water-in-fluorocarbon emulsions. *Lab Chip*. 2008,8(10):1632–9.

57. Courtois F, Olguin LF, Whyte G, Theberge AB, Huck WT, Hollfelder F, Abell C. Controlling the retention of small molecules in emulsion microdroplets for use in cell-based assays. *Anal Chem*. 2009,81(8):3008–16.
58. Gan Y, Lin Y, Guo Y, Qi X, Wang Q. Metabolic and genomic characterisation of stress-tolerant industrial *Saccharomyces cerevisiae* strains from TALENs-assisted multiplex editing. *FEMS Yeast Res*. 2018,18(5).
59. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, Philippakis AA, del Angel G, Rivas MA, Hanna M et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet*. 2011,43(5):491–8.
60. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*. 2010,20(9):1297–303.
61. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2009,25(14):1754–60.
62. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res*. 2010,38(16):e164.

Figures

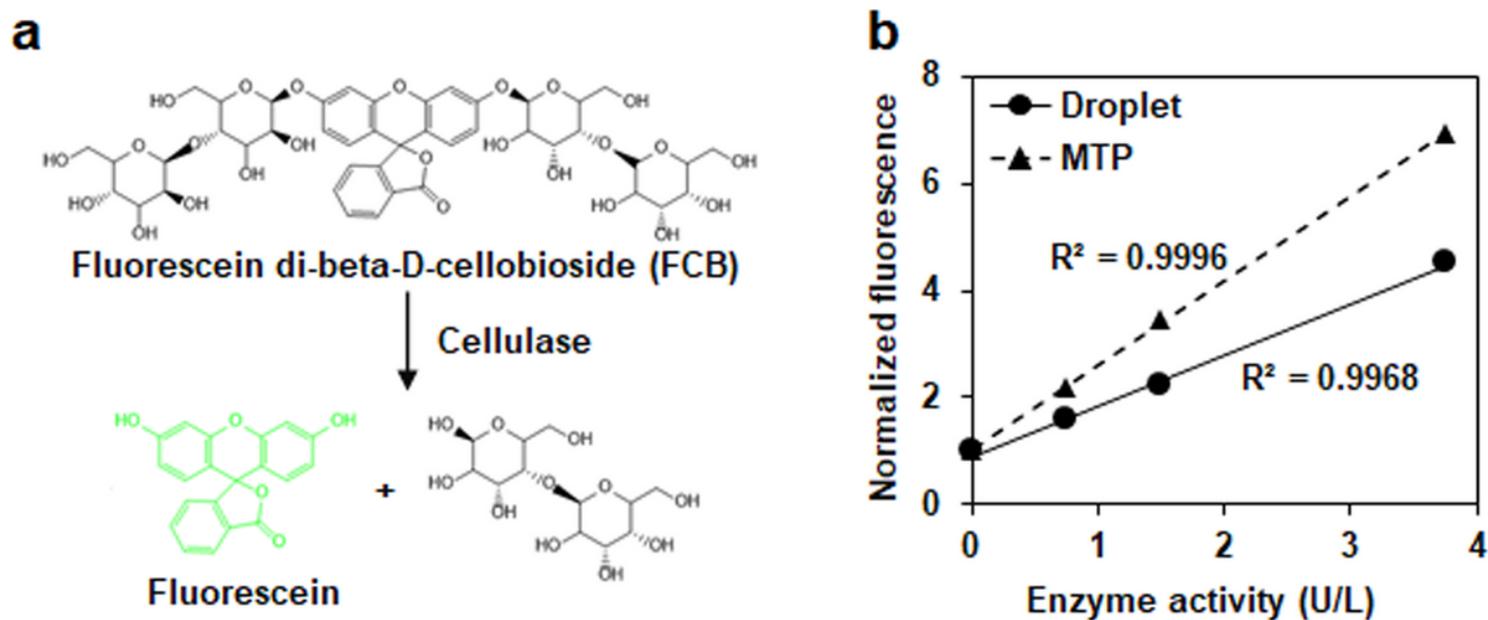


Figure 1

Cellulase activity measurement using fluorescein di-beta-D-cellobioside (FCB). a Mechanism of fluorescence-based cellulase activity measurement. Green fluorescent compound fluorescein can be released from FCB by cleavage of the β -1,4-D-glycosidic bonds under cellulase treatment. b Fluorescence-based cellulase activity assay using droplet (solid line) and MTP (dashed line) formats.

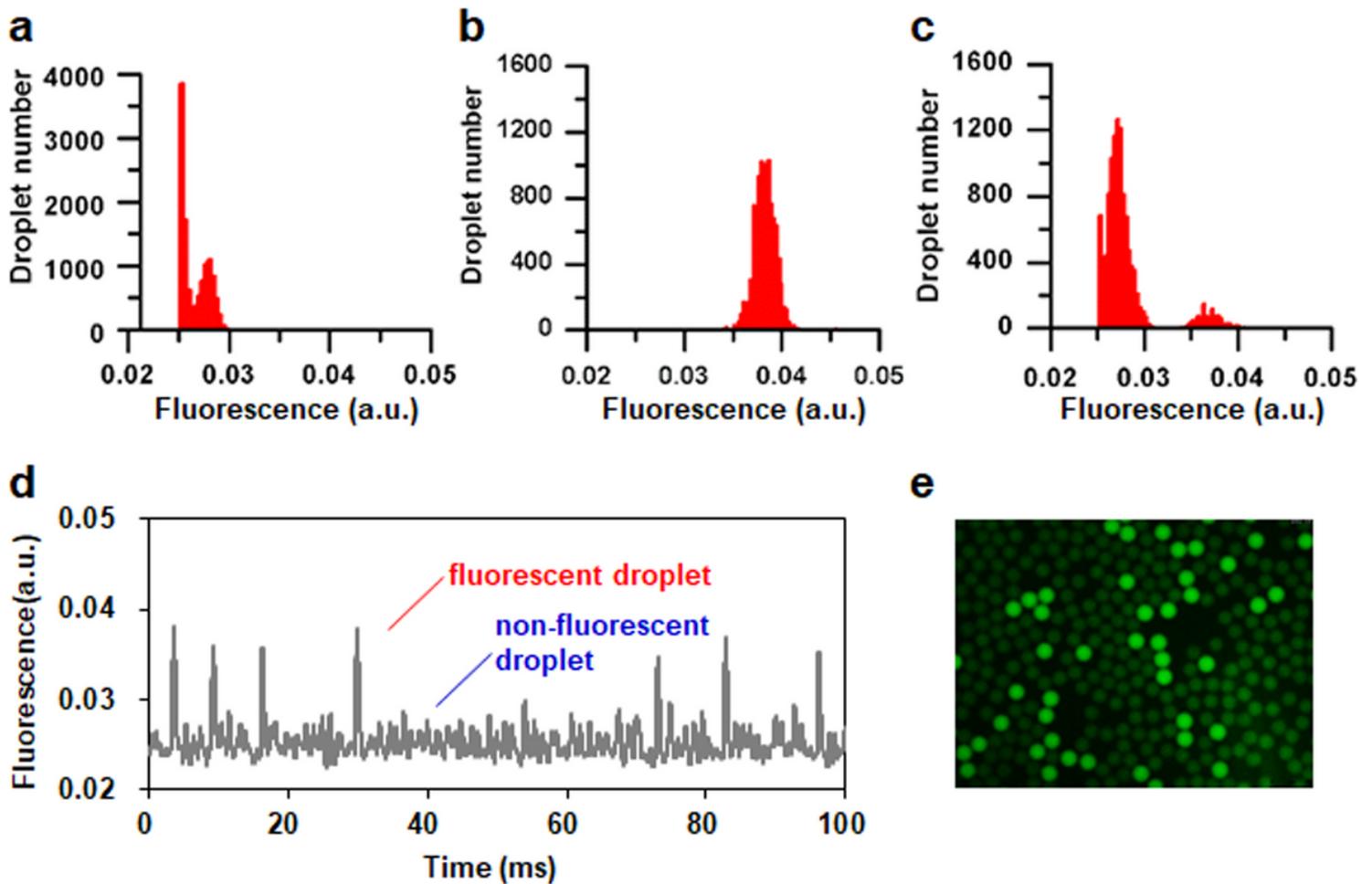


Figure 2

Effect of methanol on fluorescein leakage in droplets. a Histogram of non-fluorescent droplets only containing BMMY medium with 1% methanol. b Histogram of fluorescent droplets generated from 40 h reaction solution of cellulase in BMMY medium with 1% methanol and 20 μM fluorogenic substrate FCB. c Histogram of a binary mixture of non-fluorescent and fluorescent droplets at a 10:1 mixing ratio. d A typical recording of fluorescence signals when detecting the binary droplet mixture. e Fluorescence microscopy image of the binary droplet mixture. Droplets were incubated at 30 $^{\circ}\text{C}$ for 48 h. All histograms show approximately up to 10,000 droplets.

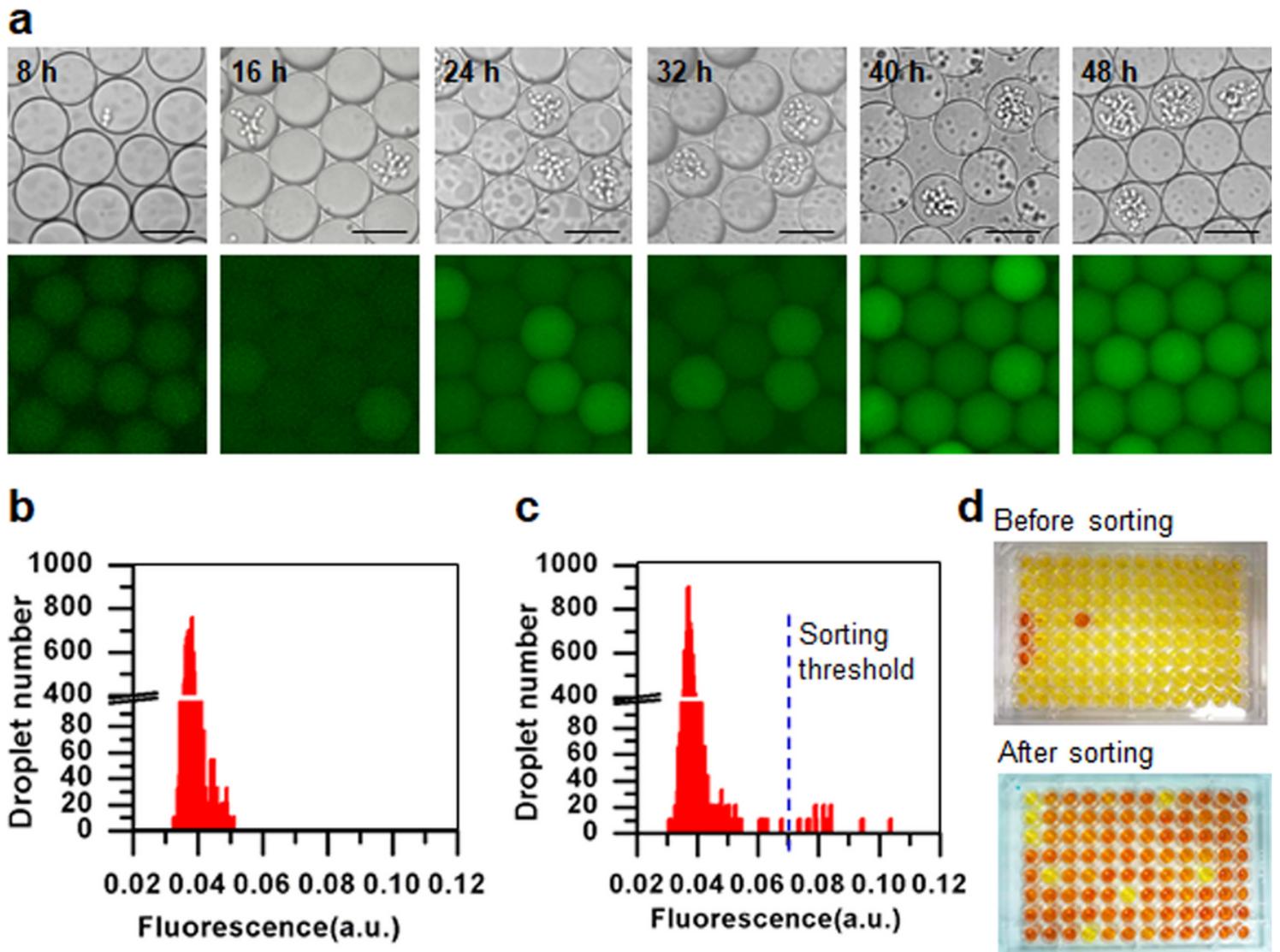


Figure 3

Establishment of the fluorescence-based droplet microfluidic screening for cellulase-producing *P. pastoris* cells. **a** Fluorescence signal detection of cellulase-producing cells of the *P. pastoris* strain SHY169 (green fluorescence-positive strain) in droplets at different time points. Scale bar: 50 μm . The exposure time was 50 ms. **b** Fluorescence histogram of droplets at 0 h. The mixture of cellulase-producing and non-cellulase producing cells at a 1:99 ratio was used for droplet generation. **c** Fluorescence histogram of droplets after incubation for 40 h at 30 $^{\circ}\text{C}$. The blue dashed line indicates the sorting threshold. **d** Cellulase activity measurement of the selected cells using the DNS method. The first three wells in the first column were negative controls of non-cellulase producing cells from the strain GS115, and the next three wells were positive controls of cellulase-producing cells from the strain SHY169.

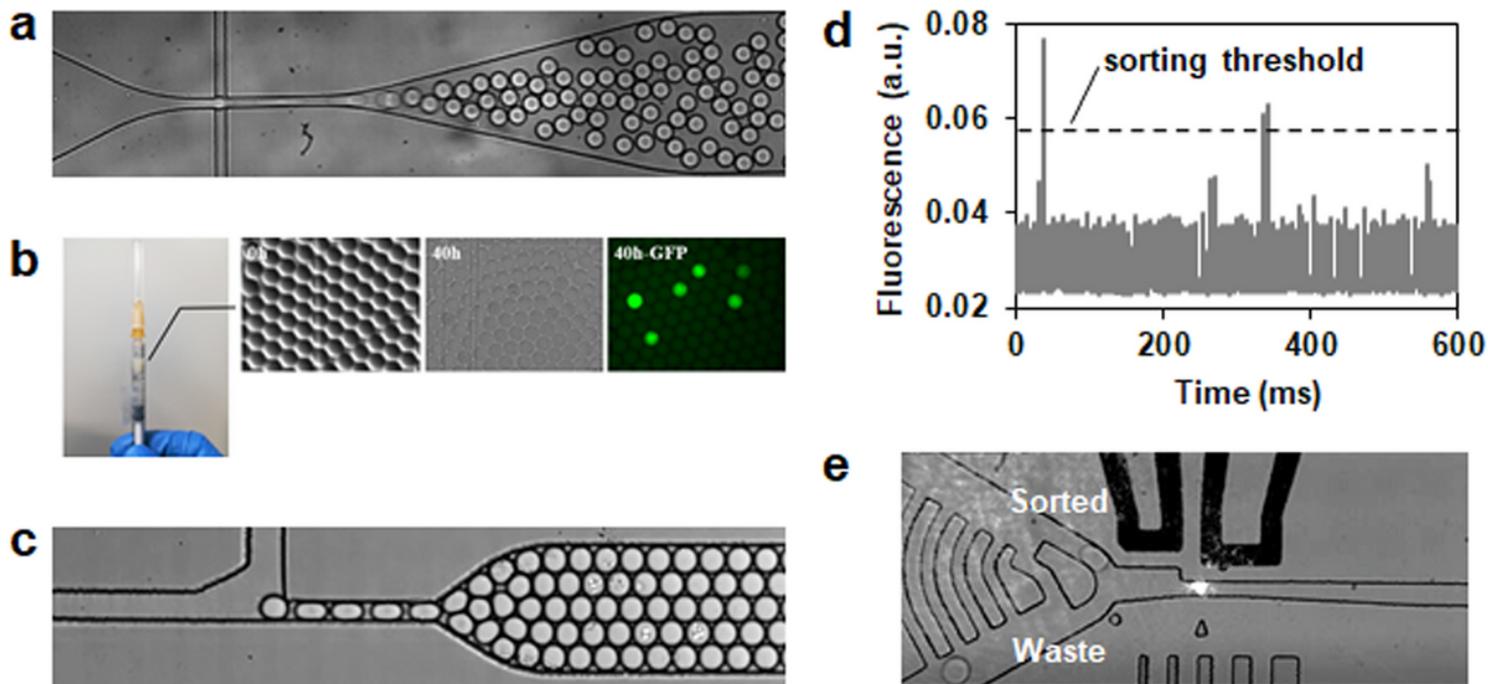


Figure 4

Workflow of the droplet-based microfluidic screening. a Droplet generation. Cells were encapsulated in microdroplets together with the substrate FCB using an appropriate cell to droplet ratio. b Droplet incubation. The emulsion was incubated off-chip in a syringe to allow cell growth, cellulase expression, and fluorescence signal development. c Droplet reinjection. After incubation, the emulsion droplets were reinjected into a sorting device. d Droplet fluorescence detection. The fluorescence signal of each droplet was detected and analyzed by the optical setup. e Droplet sorting. Droplets were sorted based on resulting fluorescence using dielectrophoresis.

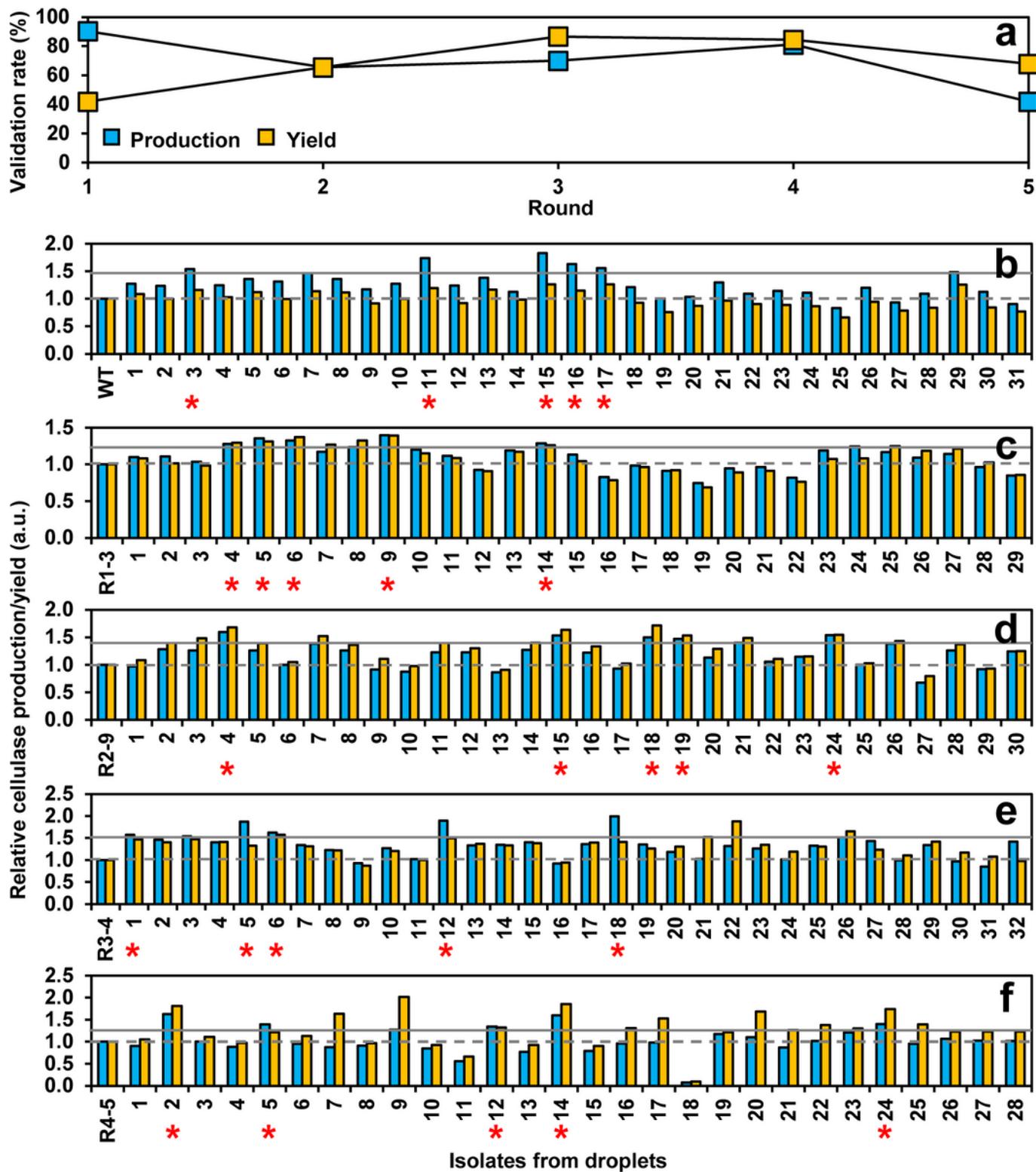


Figure 5

Primary validation of cellulase production performance by isolates from five iterative ARTP mutagenesis and screening. a Validation rates. They were calculated as the percentage of isolates with higher cellulase production and yield compared to the starting strain among the tested isolates sorted from droplets. Relative cellulase production and yield data were from Fig. 5b to 5f. b First round of mutagenesis and screening. The wild-type cellulase-producing strain SHY169 was used as a starting

strain. c Second round. d Third round. e Fourth round. f Fifth round. In c to f, the best cellulase producer was selected from the last round to be a starting strain for the next round. About thirty clones were randomly picked from the recovery plates, and their cellulase production was evaluated by flask fermentation. Production is measured in units of cellulase activity per volume and the yield is cellulase activity per 1 OD600 unit of cells. Data were normalized to the starting strain of each round. The solid and dashed lines indicate the thresholds of strain selection for production and yield, respectively. Top five cellulase hyperproducers from each round were indicated by red stars.

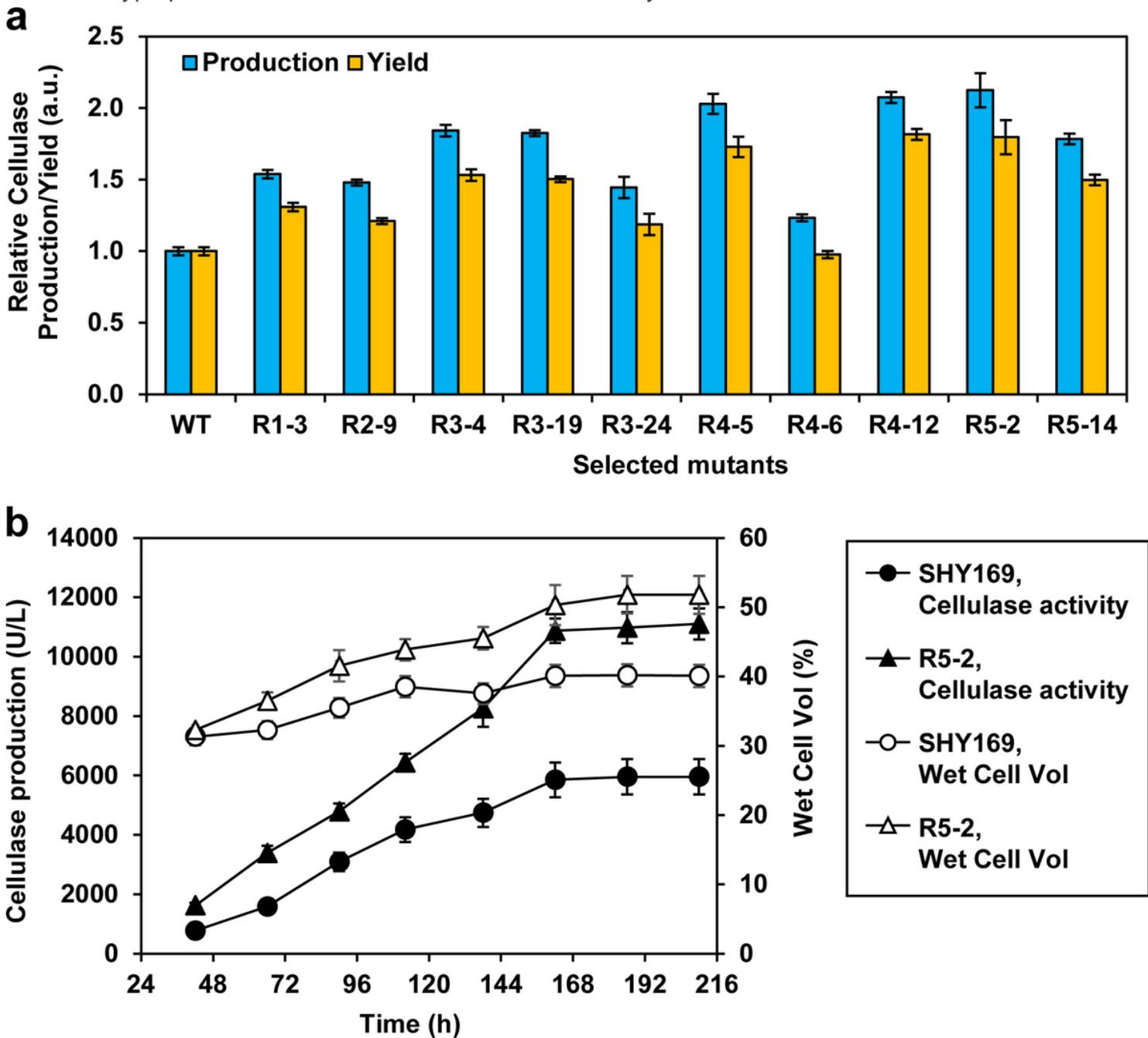


Figure 6

Enhanced cellulase production by selected mutants. a Final validation of the selected mutants for cellulase production by flask fermentation. Data were normalized to the wild-type strain SHY169. b

Comparison of fermentation performance of the best strain R5-2 and the wild-type in 5-L fermenter. Data are the average mean of triplicate samples, and the error bars represent standard deviations. Wet cell volume was measured as follows. Fermentation broth was transferred into a centrifuge tube and centrifuged at 3000× g for 10 min. Pellet volume was measured.

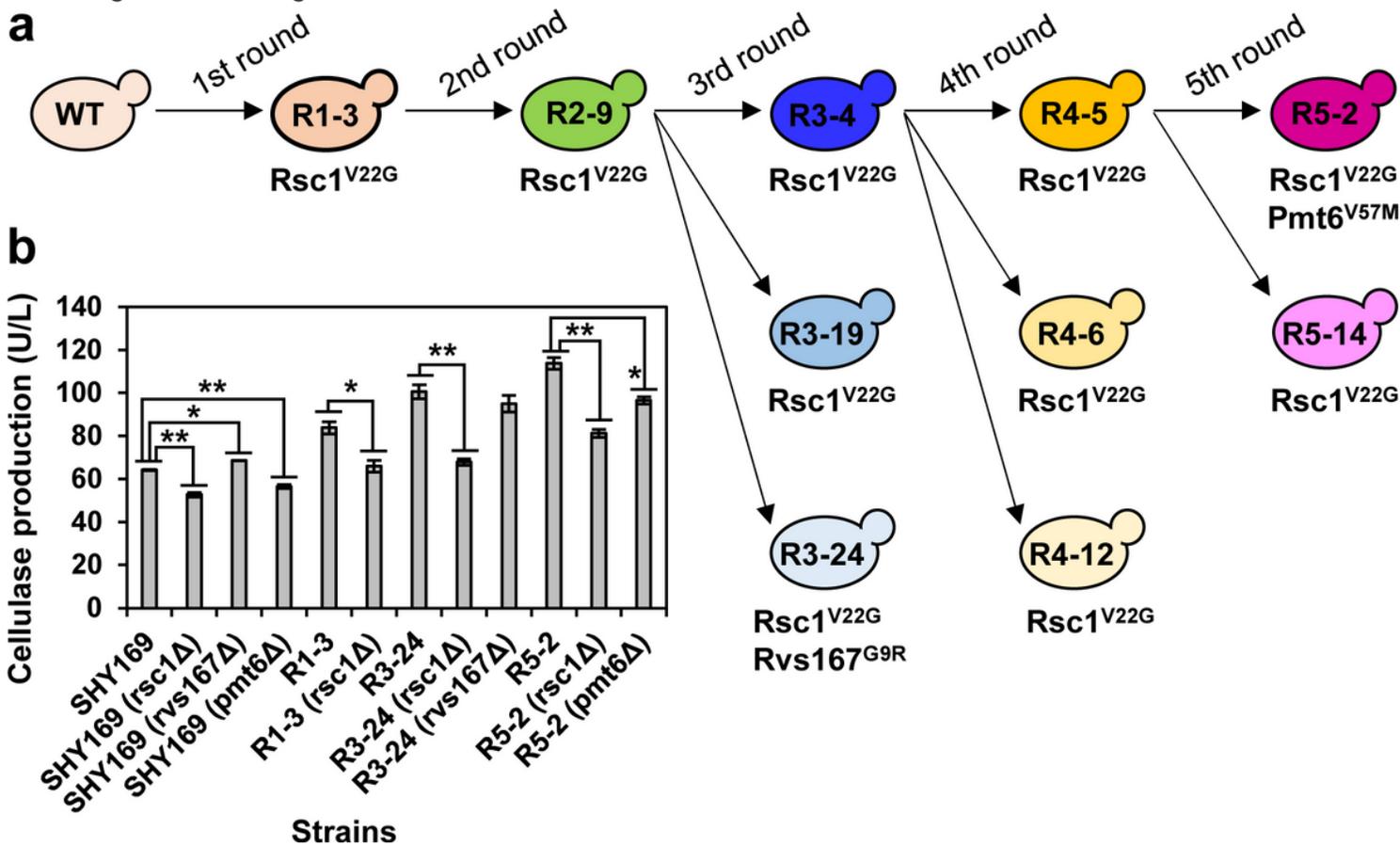


Figure 7

Identification of genomic variants using genome resequencing and gene deletion methods. a Protein point mutations uncovered by genome resequencing to be accumulated during the iterative ARTP mutagenesis and microfluidic screening. b Cellulase production in wild-type and key gene deletion strains. Data represent the mean and standard error of duplicate cultures for each strain. Statistical analysis for each group of strains including the starting strain (SHY169, R1-3, R3-24, or R5-2) and its derived deletion strains was performed using one-way ANOVA followed by Tukey's multiple-comparison posttest (**P < 0.01, *P < 0.05).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Additionalfile1TableS1S2FigureS1S2.pdf](#)
- [Additionalfile2DatasetS1.xlsx](#)

- [Additionalfile3DatasetS2.xlsx](#)