

REVIEW

Open Access



High-throughput process development from gene cloning to protein production

Manman Sun^{1,2†}, Alex Xiong Gao^{3†}, Xiuxia Liu^{1,4,5}, Yankun Yang^{1,4,5}, Rodrigo Ledesma-Amaro^{2*} and Zhonghu Bai^{1,4,5*}

Abstract

In the post-genomic era, the demand for faster and more efficient protein production has increased, both in public laboratories and industry. In addition, with the expansion of protein sequences in databases, the range of possible enzymes of interest for a given application is also increasing. Faced with peer competition, budgetary, and time constraints, companies and laboratories must find ways to develop a robust manufacturing process for recombinant protein production. In this review, we explore high-throughput technologies for recombinant protein expression and present a holistic high-throughput process development strategy that spans from genes to proteins. We discuss the challenges that come with this task, the limitations of previous studies, and future research directions.

Keywords Recombinant protein, Process development, High-throughput technology, High-throughput culture

Background

Recombinant proteins are utilized across a wide range of industries including food, chemistry, biopharmaceuticals, and biomaterials [1]. According to the latest protein expression market research report, the global protein expression market is rapidly growing, especially after the COVID-19 outbreak. The market size was valued at USD 3.18 billion in 2022 and is expected to expand at a

compound annual growth rate (CAGR) of 9.36% from 2023 to 2030 [2]. Traditionally, process development for recombinant protein production follows standard procedures established many decades ago. This process includes the selection of appropriate genetic components from existing expression toolkits through trial and error [3, 4], followed by optimizing process-related parameters one by one in shake flasks, microtiter plates (MTPs) or laboratory-scale bioreactors [5–8]. Undoubtedly, this is a time-consuming process, and it usually takes several years for a protein to develop from laboratory research to industrial production. However, the need for rapid protein production in the post-genomic era has led to the reshaping and optimization of previous process development strategies. To keep up with the demand, laboratories and enterprises are integrating high-throughput technologies into their workflows. This integration not only accelerates the production of recombinant proteins but also streamlines the entire process from gene to protein, creating a more comprehensive and time-efficient process.

[†]Co-first author.

*Correspondence:

Rodrigo Ledesma-Amaro
r.ledesma-amaro@imperial.ac.uk

Zhonghu Bai
baizhonghu@jiangnan.edu.cn

¹National Engineering Research Center of Cereal Fermentation and Food Biomanufacturing, Jiangnan University, Wuxi 214112, China

²Department of Bioengineering and Imperial College Centre for Synthetic Biology, Imperial College London, London SW7 2AZ, UK

³Division of Life Science, The Hong Kong University of Science and Technology, Hong Kong, China

⁴Key Laboratory of Industrial Biotechnology, School of Biotechnology, Ministry of Education, Jiangnan University, Wuxi 214122, China

⁵Jiangsu Provincial Research Center for Bioactive Product Processing Technology, Jiangnan University, Wuxi 214122, China



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Over the past decades, numerous methods have been developed to enable the high-throughput construction of strains [9–12]. The integration of high-throughput protein detection methods and cultivation platforms, such as MTPs, microbioreactors, and parallel fermentation systems, has made the cultivation, screening, and optimization process more efficient [13–16]. Furthermore, the adoption of novel process optimization strategies has significantly reduced the process development timeline. Herein, we review the adaptations and developments made by academic and industrial laboratories to accelerate protein production. A holistic high-throughput development strategy from genes to proteins is proposed and recommended to ensure robust and cost-effective development of the protein production processes.

High-throughput construction of strain libraries

At present, various expression systems including bacterial, yeast, insect, and mammalian cells have been developed for recombinant protein production [3, 12]. Since the target protein can have different origins and characteristics, it is challenging to predict the best expression system and still involves a significant degree of trial and error [17, 18]. Furthermore, even after selecting the expression system, choosing expression elements, such as promoters and signal peptides, is still a challenge. To maximize protein yields, it is recommended to create a library (large starting population with genetic diversity). To date, various methods have been developed for the high-throughput construction of expression strain libraries, including random mutagenesis, laboratory evolution, artificial synthesis, knockouts, and overexpression [19–22]. Among them, constructing a clone library through the combination of expression elements is the most commonly used method, which enables the systematic optimization of elements for protein expression [15, 23].

Accordingly, a given protein sequence can generate a large number of clones through the combination of promoters, signal peptides, target gene sequences, and host cells, which can be as high as $n \geq 1000$, to establish the best candidate strain [15] (Fig. 1). Over the years, various DNA assembly techniques have been developed to simplify and reduce the cost of constructing expression vectors [24]. Based on the basic principle, these methods are classified into the following three main categories: Restriction enzyme-based cloning, recombination-based cloning, and ligation-independent cloning.

Restriction enzyme-based cloning

Restriction enzyme-based cloning is a classic and widely-used method for molecular cloning that involves the digestion of DNA by restriction enzymes (RE) and the subsequent ligation of the resulting fragments. Several common systems such as the Flexi Cloning system and

Golden Gate are all based on this principle [24]. This method gained renewed attention in 2003 when Knight, T. proposed the BioBrick standard for the physical assembly of biological parts [25]. The BioBrick standard requires two special sequences for each BioBrick part, called the prefix and suffix sequences, containing REs for *EcoRI/XbaI* and *SpeI/PstI*, respectively. Among these REs, *XbaI* and the *SpeI* are isocaudomers, allowing the assembly of composite BioBrick parts [26, 27]. The key innovation of BioBrick assembly is that any two BioBrick parts can be assembled, and the resulting composite itself is also a BioBrick part that can be assembled again. Later, in 2011, Shetty et al. further developed this method and proposed the three antibiotic assembly (3 A assembly) method for the construction of BioBrick parts [27]. The 3 A assembly requires three plasmids for molecular cloning and the destination plasmid must carry a different antibiotic-resistance gene from the other two plasmids. The schematic diagram of the 3 A assembly was shown in Fig. 2. Finally, positive clones can be easily obtained through antibiotic resistance-based positive and negative selection. Compared with previous RE-based cloning, the 3 A assembly eliminates the time- and labor-intensive steps such as column cleanup and agarose gel purification during plasmid construction, increasing the throughput of molecular cloning [27–29]. This system also supports the iterative assembly of genetic components, making it an ideal tool for high-throughput construction of expression element combinations for recombinant protein production [26, 27, 30]. However, the use of 3 A assembly introduces two additional amino acids, which limits its use in scenarios with strict requirements on the protein sequence.

Recombination-based cloning

The development of recombinant cloning systems has revolutionized the construction of multiple plasmids. Among them, Gateway is probably the most successful and widely used [12]. This technology exploits a site-specific recombination system originally observed in lambda phage to transfer heterologous DNA sequences between two vectors with flanking compatible recombination attachment (*att*) sites [31]. To further improve the throughput and specificity of Gateway cloning, adjustments have been made, such as changing the sequence or length of *att* sites to clone multiple genes or fragments simultaneously [12]. However, high costs limit the widespread use of this method. To address this issue, Zhang et al. developed an alternative recombinant cloning system called SLiCE (Seamless Ligation Cloning Extract), which directly utilizes the homologous recombination activity in cell lysate prepared from the *Escherichia coli* DH10B strain expressing a lambda prophage Red/ET recombination system, enabling the assembly of multiple

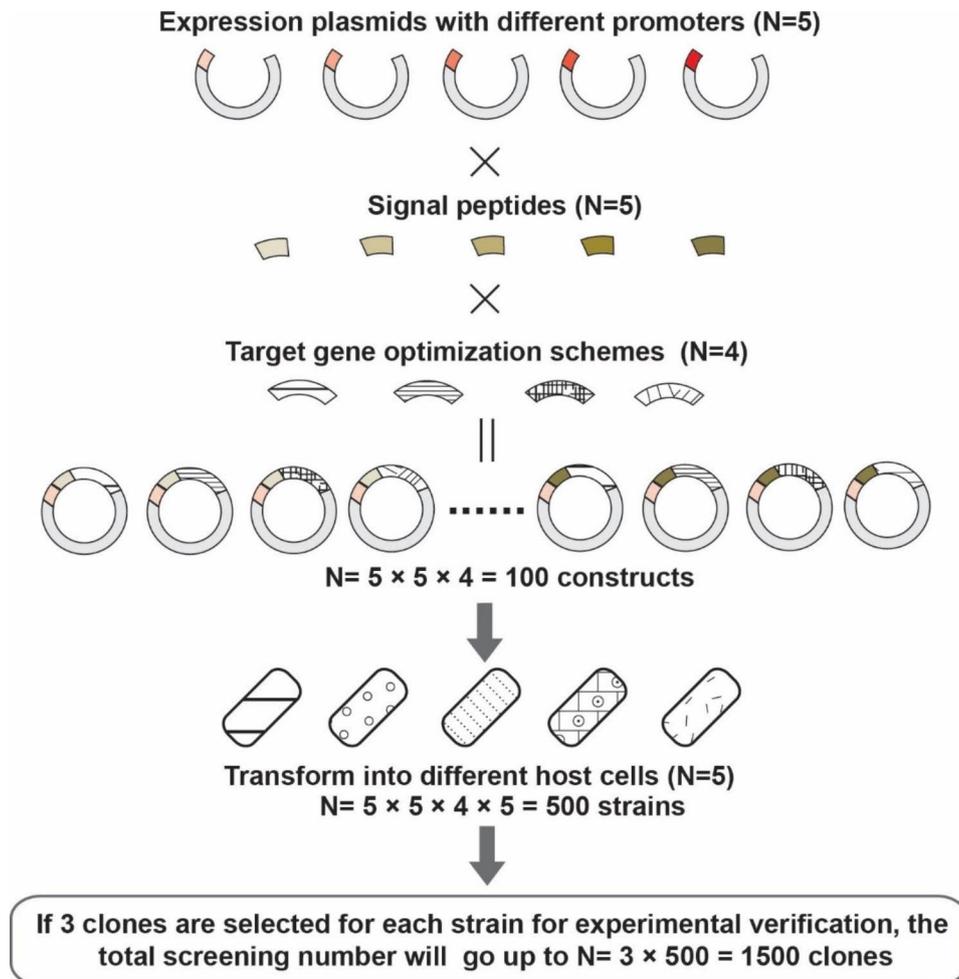


Fig. 1 Construction of recombinant strain library. A massive amount of recombinant strains is constructed according to the trial and error principle for achieving the most efficient expression combination. In this example, five promoters, five signal peptides, four target gene optimization schemes, and five phenotypes of host strain are considered for choosing the optimal clone candidate

DNA fragments into vectors in a single in vitro reaction [32]. This method was further improved by Motohashi by directly utilizing cell lysates from several common laboratory *RecA⁻E. coli* strains, including DH5 α , JM109, DH10B, XL10-gold, and Mach1 T1 [33] (Fig. 3A). Since the SliCE method does not require the use of REs and ligases, and many standard laboratory bacterial strains can serve as the source of SliCE extracts, SliCE has become a simple, efficient, and ultra-low-cost alternative to commercial kits for performing high-throughput cloning [34].

Ligation-independent cloning

Ligation-independent cloning (LIC) is a method developed 30 years ago that enables directional cloning of any fragment after generating a DNA sequence containing a single-strand (SS) complementary end [35]. The LIC method mainly uses the exonuclease activity of T4 DNA polymerase or T5 exonuclease to generate SS

complementary tails [36–38]. Here, we give a schematic for the production of recombinant DNA using T4 DNA polymerase-based LIC (Fig. 3B). As LIC does not require REs, ligases, or recombinases, it has become an inexpensive and easily adaptable method for high-throughput cloning. Hitherto, many commercially available kits based on the LIC principle have been developed, including In-Fusion from Clontech and Gibson Assembly from NEB [9]. Later, to further improve the versatility and efficiency of LIC, some improved methods have been developed, such as sequence and ligation-independent cloning (SLIC) [39], improved SLIC [40], Nicking Endonucleases based LIC (NC-LIC) [41], and uracil-excision based cloning [42, 43]. Recently, the coupling of LIC with automatic-control devices and micro-well plates has also further improved the efficiency of LIC-based plasmid construction [44, 45].

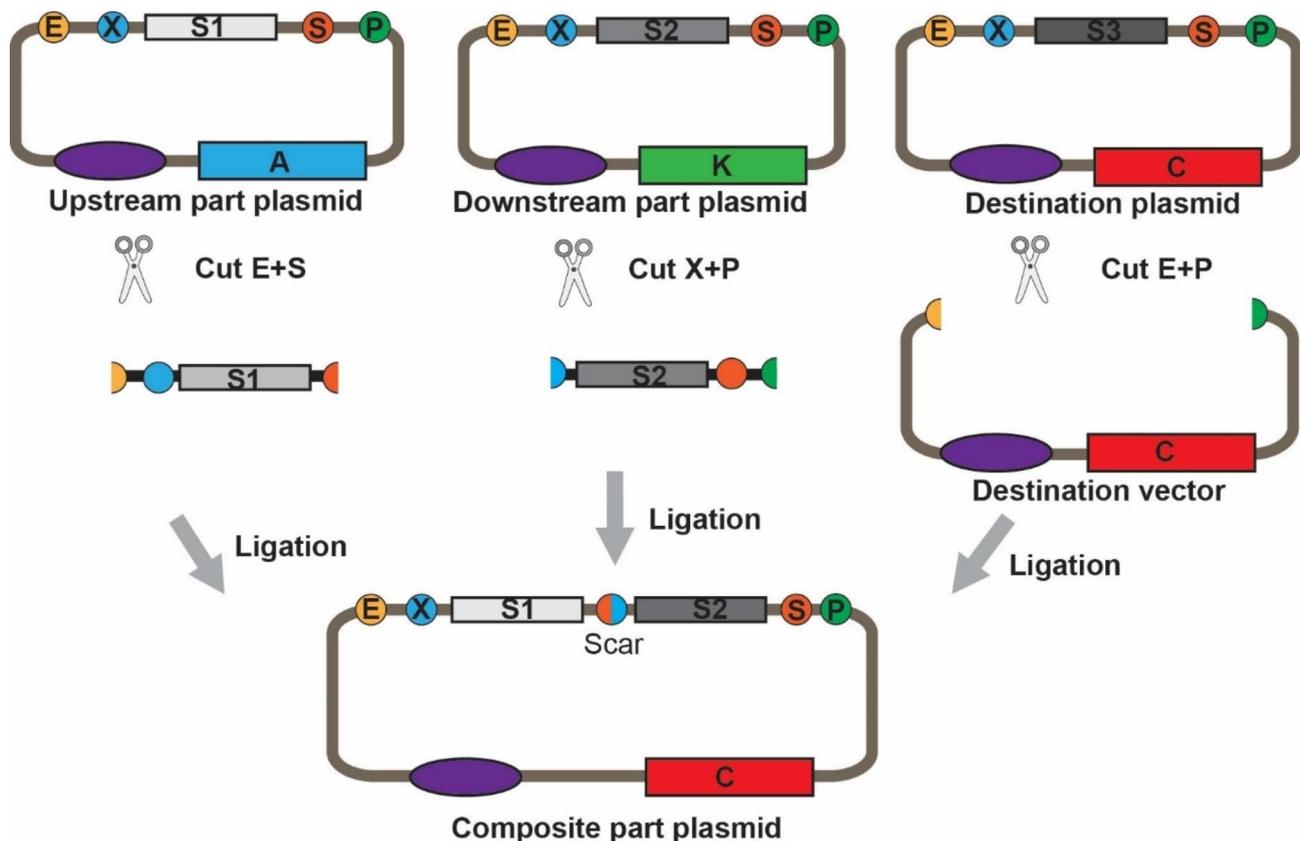


Fig. 2 Schematic diagram of the 3 A assembly process for assembling BioBrick parts. To perform 3 A assembly, the destination plasmid, the upstream part plasmid and the downstream part plasmid must have different antibiotic resistance markers from each other. Abbreviations are as follows: A=ampicillin resistance gene, K=kanamycin resistance gene, C=chloramphenicol resistance gene, S1, 2, 3=sequence 1, 2, 3, E=EcoRI, X=XbaI, S=SpeI, P=PstI. 3 A assembly works as follows: Digest the upstream part plasmid with EcoRI and SpeI. Digest the downstream part plasmid with XbaI and PstI. Digest the destination vector with the EcoRI and PstI. Then, all digested plasmids are mixed, ligated, and transformed to a solid plate supplemented with antibiotic corresponding to the destination vector resistance marker for selection. The scar represents the mixed XbaI/SpeI site

High-throughput analytical technology for protein expression

High-throughput protein expression analysis is critical in the early stages of process development. However, for many recombinant proteins, accurate quantification of their expression levels still relies on labor-intensive SDS-PAGE analysis, especially for those without specific detection methods. While several high-throughput protein analysis platforms, such as Octet™ (Pall ForteBio Corp, USA), LabChip GXII (Perkin-Elmer Inc.), and the E-PAGE™, have been developed [15], they are still in the early stage of adoption and not widely used. For most laboratories, the common method for high-throughput protein analysis is the fusion of a protein of interest with a fluorescent protein. Due to its small molecular weight and high fluorescence intensity, *Aequorea victoria* green fluorescent protein (GFP) and its mutants, such as enhanced green fluorescent protein (EGFP) and superfolder GFP (sfGFP), are widely used as fusion markers [46–48]. As shown in the work of Kovacevic and co-workers, they correlated the activity of glucose

oxidase (GOx) with GFP fluorescence [47]. In recent years, split-EGFP technology has also been developed. Instead of full-length GFP, split-EGFP technology divides the GFP into several fragments, allowing for the target protein to be fused with a small GFP fragment, thereby minimizing interference on the activity of the target protein [49]. This system has been successfully used for the high-throughput detection of a thermostable esterase Aae01 expression library (25,000 clones) in *E. coli* [50]. Another alternative strategy is the bicistronic design (BCD)-based transcriptional fusion with fluorescent proteins, where the translation of the target gene is coupled with a response gene encoding a fluorescent protein [51]. This system allows for the detection of target protein expression by monitoring fluorescence intensity, without introducing additional amino acids into the target protein [52]. Furthermore, the use of fluorescence-activating and absorption-shifting tags (FAST) emerges as another novel alternative for high-throughput protein analysis. FAST tags can be attached to target proteins, allowing for rapid, specific, and highly sensitive detection, thereby

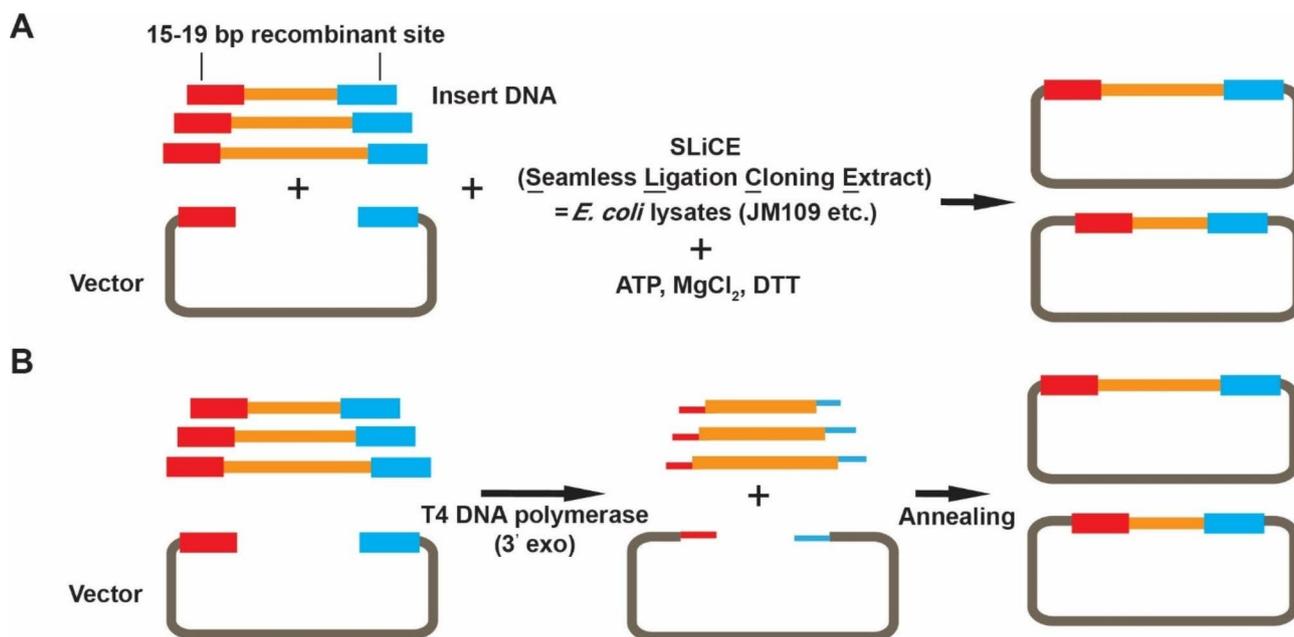


Fig. 3 Schematic diagram and the principle of the SLICE method and LIC method. **(A)** Overview of SLICE cloning. Target genes are flanked by 15–19 bp recombination sites. Laboratory *E. coli* strains' SLICE-mediated recombination between the homologous arms generates the final vector. **(B)** A schematic for the production of recombinant DNA through LIC cloning. The linearized expression vector and target gene containing complementary tails are digested by T4 DNA polymerase (3' exo) and then transformed into *E. coli* for in vivo ligation after annealing

enhancing the efficiency and precision of high-throughput protein expression analysis [53].

Although the fusion expression strategy is convenient and accurate, it still has limitations in detecting transiently expressed or fast-degrading proteins with short half-lives [54]. To overcome this issue, biosensors such as the STEP sensor (sensor for transiently expressed proteins) have been developed, providing a solution for the high-throughput detection of protein expression, especially for transiently expressed proteins [55].

High-throughput cultivation platform

The development of a reliable and cost-effective high-throughput cultivation platform is crucial due to its time-consuming and costly nature. Over the past decades, various miniaturized culture devices have been developed, enabling cultivation at milliliter, microliter, or even picoliter scales [14, 56]. A good example of these devices is the microfluidic-based cultivation system [56]. To date, various microfluidic bioreactors, single-cell habitats, trapping cavities, and cultivation chambers have been developed. Based on the cells' degree of freedom, these microfluidic culture devices can be classified into different dimensions (ranging from 0-dimensional (0D) to 3D) [14, 56] (Fig. 4A). Despite their advantages, microfluidic-based cultivation systems increase the risk of contamination due to the use of continuous single-phase flow [57]. To overcome this problem, researchers further developed microdroplet technology [58, 59] (Fig. 4B). By separating

the carrier fluid from the culture medium and encapsulating microbial cells in droplets, microdroplet technology eliminates contamination [59]. These innovative microdroplet systems have been successfully used for microorganism enrichment [60], high-throughput characterization and screening of strains [19, 61], adaptive evolution [59], etc. Another representative cultivation platform is the microliter-level microtiter plates (MTPs) system, the advantages of high-throughput, easy-to-operate, and low-cost advantages of MTPs make it a widely used cultivation platform [15, 62, 63]. To date, various MTPs formats (6–1,536 wells) have been developed and many auxiliary devices such as pipetting robots, autosamplers, and microplate readers have been made compatible with MTPs [63]. MTPs have now become a cheap alternative to shake flasks for strain cultivation.

To better meet the needs of high-throughput process development of protein production, a vast number of miniaturized bioreactors, such as the miniature 10-ml stirred-tank bioreactor, 10-mL scale microbioreactor, 5-ml Applikon microreactor, 3-ml Biocurve, and μ -bioreactor system BioLector have also been developed [62], and a few of these have been successfully commercialized. In Table 1, we listed and compared several commercial mini-bioreactor systems. Depending on the culture broth mixing mechanism, they can be categorized as bubble column- or microplate-based mini-bioreactors and stirred mini-tank bioreactors [15]. These platforms facilitate strain screening: They not only allow

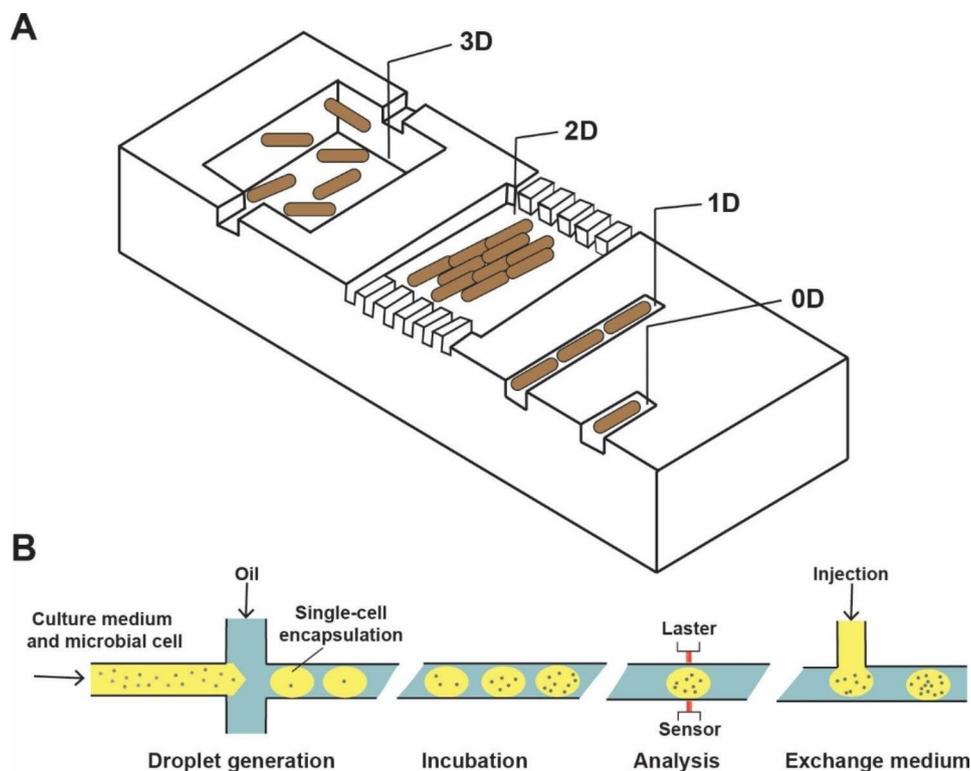


Fig. 4 Microfluidic-based micro-cultivation system. **(A)** Overview of the geometric principles of microbial single-cell reactors. Nanoliter chambers for 3D cultivation, picoliter 2D chambers to hold cell monolayers, and femtoliter channels for 1D linear single-cell rows and single-cell traps. **(B)** Droplet-based microfluidic micro-cultivation system

high-throughput screening of strains under controlled conditions, but also some of them can collect a wealth of process information online (protein titer, protein quality, specific productivity, cell fitness, or robustness) for each clone. However, these miniaturized bioreactors were considered not ideal for high-throughput optimization due to the lack of independent feeding systems and the inability to achieve high cell densities. Recent advances in some mini-reactors such as the microbial Ambr15f fermentation system (Ambr15f) have overcome this limitation. The pumped liquid line in Ambr15f can feed each vessel as needed, making it a good scale-down model for fermentation parameter optimization [64].

In addition to the above miniaturized bioreactors, researchers have also developed a variety of parallel fermentation systems, such as the microbial Ambr250 system [65], the 4×1 L Biocuber system [23], Multifors 2 system [66], BioXplorer system [67], multi-bioreactor system BIOSTAT® Qplus, and Dasgip parallel bioreactor system [68, 69] (Table 2). These systems consist of multiple small stirred bioreactors that operated simultaneously, they offer several advantages over traditional single bioreactor setups, including increased throughput and reduced resource consumption, and reduced space

requirements [62]. With larger culture volumes (usually 50–1000 mL) than micro-bioreactors (<15 mL), parallel fermentation systems enable us to perform a complete downstream analysis. These systems have now been widely used in fermentation parameter optimization [15, 23].

High-throughput strategy for process optimization

Optimization of the medium composition

After a production strain has been chosen, it becomes necessary to further optimize the components of the cultivation medium. Traditional one-factor-at-a-time (OFAT) methods in shake flasks are low-throughput and fail to consider interactions between medium components. To mitigate this limitation, a combination of high-throughput cultivation platforms such as micro-bioreactors or MTPs with experimental design (DoE) can be utilized [23, 70–72]. DoE enables the study of interactions between variables and reliable prediction of results in unexplored conditions [73, 74], which can significantly reduce the number of necessary experiments and increase experimental throughput. Moreover, model predictive control (MPC) can also be integrated into this kind of DoE-based framework, allowing the precise,

Table 1 Comparison of several commercial mini-bioreactor systems and their applications in modern bioprocess development/research

Name	BioLector	Micro-24	Micro-matrix	2mag bioreactor 48	Ambr15f
Characteristics	Microplate-based bioreactor; Operates with non-invasive optical sensors; Real time culture monitoring; Can be upgraded to fully automated units	Bubble column-based bioreactor; 24 simultaneous cultures with independent control of each reactor; individual pH, DO, and temperature sensors ; Real time monitoring	Microplate-based mini-bioreactor; 24 simultaneous cultures with independent control of each reactor; Individual liquid addition; Real time monitoring	Microplate-based mini-bioreactor; 48 parallelized mini reactors; Non-invasive real-time measurement of pH and DO	Micro-stirred tank bioreactor; Single-use pH and DO sensors; Available with temperature compensation for low-temperature application; Real time monitoring
Working volume	0.8–1.5 mL/well	3–7 mL/column	1–7 mL/well	8–15 mL/well	10–15 mL/vessel
Availability of feed	Require an external liquid handling station	Integrated liquid feed individually per column	Integrated liquid feed individually per bioreactor	Require the integration of a pipetting robot	Integrated liquid feed individually per vessel
Application areas	Cell and strain screening; Growth characterization analysis; Medium screening and optimization; Culture condition optimization; Systematic biology; Gene/proteomics studies	Clone screening; Optimizing temperature, pH, and DO conditions for cell growth; Optimizing the induction condition for protein expression; Optimizing medium composition; Process development studies	Screening cell-line, microbial, and/or yeast libraries; Process development studies; Process optimization studies; Small volume cultivations	Clone selection; Media screening and optimization; Optimization of the process design; Studies of gene and protein expression	Clone selection; Media and feed optimization; Process intensification; Development of advanced cell therapies
Reference	https://doi.org/10.1186/1472-6750-11-22 https://doi.org/10.1186/1475-2859-8-31	https://doi.org/10.1002/btpr.522 https://doi.org/10.1002/bit.22031	https://doi.org/10.2144/btn-2019-0063	https://doi.org/10.1007/s00449-022-02798-6	https://doi.org/10.1002/btpr.2534
Website of suppliers	https://www.m2p-labs.com/	http://www.pall.com/main/biopharmaceuticals	http://www.applikon-bio.com	https://www.2mag.de/en/produkte-e/bioreactor-e/bioreactor-48-e.html	http://www.tapbiosystems.com

real-time adjustment of cultivation conditions based on predictive models. This comprehensive approach takes into account the interdependencies between different medium components and process parameters, thus facilitating the efficient determination of the optimal medium composition [75, 76].

Optimization of the fermentation parameters

Fermentation parameters such as dissolved oxygen (DO), feeding rate, pH, and agitation speed significantly affect protein expression. Considering the importance of these parameters at the industrial scale, it is essential to optimize them in stirred bioreactors [77]. Traditional optimization methods based on iterative experiments in laboratory-scale bioreactors are expensive and time-consuming. Therefore, stirred mini-tank bioreactors or parallel fermentation systems combined with the DoE strategy are increasingly used. For instance, Janakiraman

et al. optimized monoclonal antibody production in CHO cells by using the Ambr system and DoE to determine the optimal growth temperature, production temperature, and pH [78]. Our previous study optimized the fermentation parameters for OmlA antigen production by combining the Biocuber system and DoE [23]. Recent advancements in high-throughput stirred bioreactors, such as the Ambr 15f system software's compatibility with DoE packages and the BioPAT® MFCS/win module's facilitation of automated optimization experiments, have greatly simplified the DoE-based process optimization, enhancing reliability and reproducibility [64, 79]. Successful applications include the optimization of malaria vaccine production in *Pichia pastoris* [79].

Building on these innovations, MPC can further optimize the process by enabling dynamic, real-time adjustments of fermentation parameters [75, 76]. Simultaneously, a comprehensive digital infrastructure

Table 2 Summary of basic characteristics of several commercial parallel fermentation systems

Name	Ambr250 system	Biocuber system	Multifors 2
Characteristics	Contain 12 or 24 fully featured single-use mini bioreactors; Individual control for each bioreactor vessel; Integrated liquid feed individually per bioreactor; Real time culture monitoring	Contain 4 independently controlled bioreactors; Integrated liquid feed individually per bioreactor; Real time culture monitoring	Contain 6 independently controlled bioreactors; Integrated liquid feed individually per bioreactor; Real time culture monitoring
Working volume	100–250 mL	300 mL to 1 L	150 mL to 1 L
Application areas	Applicable for process development; Process optimization; Scale-down studies; Cell culture and microbial fermentation	Applicable for the process development of bacteria, yeast, mammalian cell cultures; Clone screening; DoE optimization	Applicable for the process development of microorganisms and mammalian cell cultures; Clone screening; DoE optimization
Reference	https://doi.org/10.1002/biot.201700766	https://doi.org/10.1007/s00253-022-11918-x	https://doi.org/10.3389/fbioe.2021.695306
Website of suppliers	https://www.sartorius-stedim-tap.com/tap/cell_culture/ambr_250.htm	http://www.bioyd.com/	https://www.infors-ht.com/en/bioreactors/bench-top-bioreactors/multifors2/
Name	BIOSTAT® Qplus multi-bioreactor system	BioXplorer 100/400	Dasgip parallel bioreactor system
Characteristics	Fully independent control of up to 12 expandable culture vessels; Specifically designed for early process development and multivariable process optimization; Offers a broad range of measurement and automation features	Parallel processing of 4 or 8 independent reactors; Integrated individual liquid and gas feeds with options for gas mixing; Real time culture monitoring	Fully independent control of up to 16 expandable culture vessels; All control systems can be accessed remotely by one or more remote operators simultaneously
Working volume	150 mL to 1 L	20 to 100 mL; 20 ml to 400 ml	1.3 to 4.3 L
Application areas	Applicable for microbial, mammalian, insect and plant cell growth studies; Clone screening; DoE optimization; Small-scale expression of recombinant proteins	Microbial fermentation; C1 Gas fermentation and cell culture; Process development, DoE optimization	Applicable for process development of bacteria, yeast, fungi, and mammalian cell cultures; DoE optimization
Reference	https://doi.org/10.1038/nmeth.f.340	https://doi.org/10.3389/fbioe.2021.695306	https://doi.org/10.1186/1472-6750-12-96
Website of suppliers	https://www.sartorius.com/en	https://helgroup.com/products/bioreactors/	http://www.dasgip.com

enhances data management, analytics, and automation, leading to a more efficient and reliable DoE-based fermentation process optimization [80].

Current limitations

Although high-throughput process development has shown satisfactory results, there are still some limitations that need to be addressed. Firstly, there is a lack of systematic integration of high-throughput technology to cover the entire process of protein production. Previous applications of high-throughput technology have mainly focused on a small part of the overall protein production process [81]. Although several researchers have begun to use high-throughput technology in an integrated manner, the practice still needs to be more widespread to cover the whole process [23, 82, 83]. Secondly, most previous studies have paid little attention to the reliable scale-up of the fermentation process. They either finish after establishing the fermentation process in miniaturized bioreactors or just copy the fermentation parameters settings established in these scale-down bioreactors to large-scale ones [84, 85]. Microbial phenotypic heterogeneity may

be aggravated during the scale-up process due to the formation of gradients such as oxygen and substrates in large-volume vessels, which have also been overlooked and not thoroughly discussed [86]. Finally, process development strategies need to be adjusted according to the existing high-throughput platforms. For instance, most production strains in previous studies were screened solely based on protein titers in uncontrolled cultivation platforms [19, 20]. With the development of micro-bioreactors with online monitoring systems, multi rounds of screening with multiple performance criteria evaluations (protein titer, protein quality, specific productivity, cell fitness, or robustness) should be adopted. Additionally, as the throughput of the culture platform increases, the combination of experimental variables and experimental design should be carefully explored to obtain the best results [87].

High throughput process development for recombinant protein production

Here, we summarized a representative development process for recombinant protein production. The holistic

process includes high-throughput clone construction and screening, high-throughput production process optimization, and reliable scale-up of the production process (Fig. 5).

High-throughput screening

Constructing expression libraries is widely utilized to identify the optimal expression element combination for the production of a given protein. As described above, various high-throughput cloning methods have been developed, allowing us to construct large libraries of strains in a short time. For the screening of the production strain from a large pool of candidates ($n \approx 1000$), a multi-round of screening strategy is highly beneficial [13, 15, 88, 89] (Fig. 5). The first round of screening is typically performed in an uncontrolled MTPs system to select the most promising clones based on selection criteria such as protein titer (g/L), which would reduce the number of clones for the following round ($n \approx 50-100$). At this stage, the glucose limited fed-batch technology, a strategic maneuver that optimizes the fermentation process by precisely controlling the supply of glucose, can be adopted, to fully unleash the production potential of clones [90]. The second round of screening can be conducted in a highly parallel controlled platform, such as micro-24 or micro-matrix, further reducing the number of candidate clones ($n \leq 10$). Finally, the performance of

these selected strains is comprehensively evaluated in cultivation systems closer to actual production, such as stirred micro-bioreactor and parallel fermentation system, to identify the best production strain. It is worth noting that the strain ranking in the parallel fermentation system and micro-bioreactor may differ from that in the MTPs, reflecting the impact of environmental factors on microbial performance [23, 89].

High-throughput production process optimization

Once the production strain has been selected, further experimental optimization is needed to determine the optimal production medium and fermentation parameters for the production of the target protein. Considering a large number of experimental variables, a high-throughput cultivation platform combined with the DoE strategy, is strongly recommended at this stage [71, 91] (Fig. 5). In this context, the integration of MPC and a comprehensive digital infrastructure can further enhance the optimization process. Following this strategy, the experimental design optimization of medium components is suggested to be performed in MTPs, while the optimization of fermentation parameters is preferably performed in stirred micro-bioreactors or parallel fermentation systems [23]. These combined high-throughput optimization strategies, supported by MPC and a comprehensive digital infrastructure, allow us to quickly

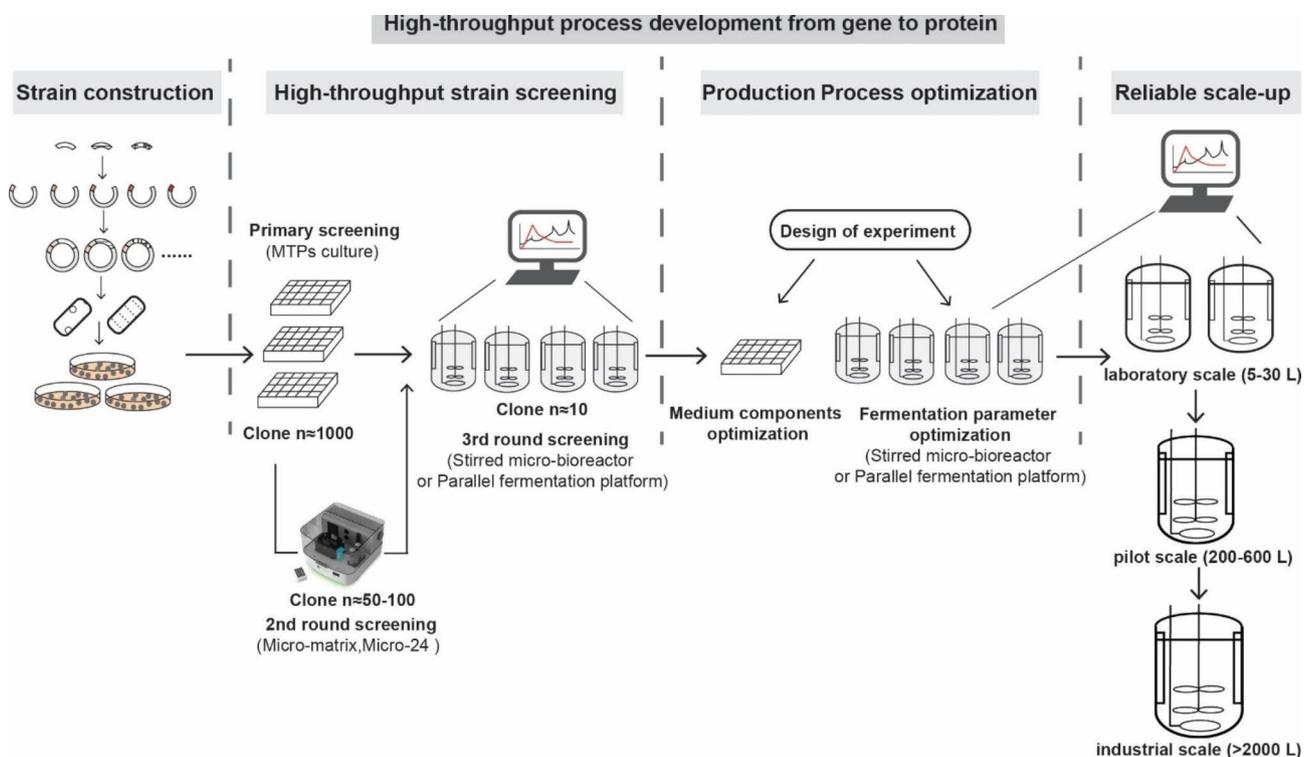


Fig. 5 Representative flow chart of high-throughput process development from gene cloning to protein production in the current bio-industry sector. Model predictive control (MPC) and a comprehensive digital infrastructure can be integrated to accelerate this process

identify the critical process variable and determine their “design space” [15, 23, 92]. Ultimately, the optimal setting level for the key variable can be easily determined through a limited number of experiments.

Reliable scale-up of the production process

The fermentation process established in the scale-down model needs to be scaled up to a larger scale bioreactor for further evaluation or actual commercial production (Fig. 5). To ensure the reliability of the scale-up, it is important to adopt a suitable scale-up criterion. Constant oxygen mass transfer coefficient ($k_L a$), constant specific power input (P/V), constant impeller tip speed, and constant dissolved oxygen concentration are four commonly used scale-up criteria in the fermentation industry [93]. However, due to the complexity of the cell culture process and the varied characteristics of recombinant proteins, it seems that no criterion can be universally applied with a high success rate [94]. The actual selection of scale-up criteria should be based on the specific characteristics of the fermentation process. Generally, a constant $k_L a$ strategy is recommended for scaling up aerobic microorganisms [23]. Whereas constant P/V is often used as a scale-up criterion for early industrial penicillin fermentation and low-energy input fermentation [95], and this strategy is limited in fermentation processes that require high-energy input, such as the recombinant *E. coli* culture [96]. Constant tip speed is ideal for scaling up antibiotic fermentation and evaluating the possibility of hyphal rupture in the fermentation of branched yeast, filamentous bacteria, and fungi [97], but it is less useful for single-cell fermentation. When heat transfer is a limiting factor for fermentation scale-up, such as high-density fermentation of *Pichia pastoris* using methanol as a carbon source, scale-up based on constant dissolved oxygen concentration is preferred [98]. Additionally, strain and inducer modifications, cell physiology manipulations, and bolus feeding with pulses strategy can also be adopted to reduce cell phenotypic heterogeneity during the scale-up process [86]. Finally, it is worth noting that, the experimenter’s intuition and expertise are also crucial in the scale-up process [99].

Conclusion and future outlook

In this article, we reviewed the high-throughput technologies that have been developed and applied to the recombinant expression of proteins. We also proposed a holistic high-throughput process development strategy. To further accelerate the process development for protein production, there is still much work to be done. A primary imperative is the broad-scale integration of automated laboratory processes. This integration, achieved by harmonizing automated sample preparation with cultivation platforms and aligning them with high-throughput

analytical tools, serves to reduce human error while enhancing laboratory efficiency [100, 101]. Connecting these disparate elements of the protein production workflow allows us to create a unified, efficient, and predominantly automated process that could redefine the standards of future protein production. Potential areas for future developments lie in refining detection tools and experimental equipment. For instance, biosensors, a staple in metabolic engineering, could be further optimized for protein expression detection, especially in the case of secreted proteins. The introduction of disposable, pre-sterilized bioreactors could mitigate sample contamination and decrease labor-intensive preparation. In addition, as the application of various high-throughput technologies increases, the generation of data multiplies and solid systems to manage, store and analyze the obtained results need to be developed. A comprehensive digital infrastructure can be established for managing, sharing, and analyzing experimental data throughout the development process [80]. Moreover, the creation of a public database of protein expression data and the introduction of bioinformatics analysis are also necessary. The public database can collect and share the conditions and results of protein expression from laboratories worldwide. By inferring rules based on bioinformatics analysis of previous data, researchers can predict which expression elements and culture conditions may be successfully used for protein expression with specific characteristics, thus greatly reducing the workload of clone construction and process optimization [102]. Finally, it is expected that new artificial intelligence (AI) and machine learning (ML) techniques will play a critical role in such development. They can improve process efficiency, enhance product quality, and reduce production costs. Additionally, AI and ML can support scale-up, data integration and visualization, and automation, leading to faster and more cost-effective production of bio-products [103, 104].

Author contributions

M.S. wrote this manuscript with assistance from X.G. and X. L. R.L.-A., Y.Y., and Z.B. helped revise this manuscript.

Funding

This work was supported by National Natural Science Foundation of China (No.22078128, 21938004), the 111 Project (No. 111-2-06), the Fundamental Research Funds for the Central Universities (No. JUSRP221032), Postgraduate Research & Practice Innovation Program of Jiangsu Province KYCX21-2028, National First-Class Discipline Program of Light Industry Technology and Engineering (LITE2018-24), British Council 527429894 and BBSRC BB/T013176/1.

Data Availability

Not applicable.

Declarations

Competing interests

The authors declare no competing interests.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Received: 24 May 2023 / Accepted: 19 August 2023

Published online: 15 September 2023

References

- Wang Y, Katyal P, Montclare JK. Protein-engineered functional materials. *Adv Healthc Mater.* 2019;8:e1801374.
- <https://www.grandviewresearch.com/industry-analysis/protein-expression-market>. Accessed 28 July 2023.
- Tripathi NK, Shrivastava A. Recent developments in bioprocessing of recombinant proteins: expression hosts and process development. *Front Bioeng Biotechnol.* 2019;7:420.
- Teh MY, Ooi KH, Danny Teo SX, Bin Mansoor ME, Shaun Lim WZ, Tan MH. An expanded synthetic biology toolkit for gene expression control in acetobacteraceae. *ACS Synth Biol.* 2019;8:708–23.
- Liu Y-K, Lu C-W, Chang J-Y, Lu C-F, Tan C-C, Huang L-F. Optimization of the culture medium for recombinant protein production under the control of the α My3 promoter in a rice suspension-cultured cell expression system. *Plant Cell Tissue Organ Cult.* 2018;132:383–91.
- Hevekerl A, Kuenz A, Vorlop K-D. Filamentous fungi in microtiter plates—an easy way to optimize itaconic acid production with *aspergillus terreus*. *Appl Microbiol Biotechnol.* 2014;98:6983–9.
- Ritacco FV, Wu Y, Khetan A. Cell culture media for recombinant protein expression in chinese hamster ovary (CHO) cells: history, key components, and optimization strategies. *Biotechnol Prog.* 2018;34:1407–26.
- Sohoni SV, Nelapati D, Sathes S, Javadekar-Subhedar V, Gaikawari RP, Wangikar PP. Optimization of high cell density fermentation process for recombinant nitrilase production in *E. coli*. *Bioresour Technol.* 2015;188:202–8.
- Jia B, Jeon CO. High-throughput recombinant protein expression in *Escherichia coli*: current status and future perspectives. *Open Biol.* 2016;6:160196.
- Mason DM, Weber CR, Parola C, Meng SM, Greiff V, Kelton WJ, et al. High-throughput antibody engineering in mammalian cells by CRISPR/Cas9-mediated homology-directed mutagenesis. *Nucleic Acids Res.* 2018;46:7436–49.
- Liao B, Chen X, Zhou X, Zhou Y, Shi Y, Ye X, et al. Applications of CRISPR/Cas gene-editing technology in yeast and fungi. *Arch Microbiol.* 2022;204:79.
- Hunt I. From gene to protein: a review of new and enabling technologies for multi-parallel protein expression. *Protein Expr Purif.* 2005;40:1–22.
- Bareither R, Bargh N, Oakeshott R, Watts K, Pollard D. Automated disposable small scale reactor for high throughput bioprocess development: a proof of concept study. *Biotechnol Bioeng.* 2013;110:3126–38.
- Ladner T, Grünberger A, Probst C, Kohlheyer D, Büchs J, Delvigne F. Application of mini- and micro-bioreactors for microbial bioprocesses. *Curr Dev Biotechnol Bioeng.* 2017;433–61.
- Long Q, Liu X, Yang Y, Li L, Harvey L, McNeil B, et al. The development and application of high throughput cultivation technology in bioprocess development. *J Biotechnol.* 2014;192:323–38.
- Moses S, Manahan M, Ambrogely A, Ling WLW. Assessment of AMBR™ as a model for high-throughput cell culture process development strategy. *Adv Biosci Biotechnol.* 2012;3:918–27.
- Makino T, Skretas G, Georgiou G. Strain engineering for improved expression of recombinant proteins in bacteria. *Microb Cell Fact.* 2011;10:32.
- Retallack DM, Jin H, Chew L. Reliable protein production in a *Pseudomonas fluorescens* expression system. *Protein Expr Purif.* 2012;81:157–65.
- Zhang G, Chen Y, Li Q, Zhou J, Li J, Du G. Growth-coupled evolution and high-throughput screening assisted rapid enhancement for amylase-producing *Bacillus licheniformis*. *Bioresour Technol.* 2021;337:125467.
- Meng L, Gao X, Liu X, Sun M, Yan H, Li A, et al. Enhancement of heterologous protein production in *Corynebacterium glutamicum* via atmospheric and room temperature plasma mutagenesis and high-throughput screening. *J Biotechnol.* 2021;339:22–31.
- Yamada O, Nishibori N, Hayashi R, Arima T, Mizutani O. Construction of transcription factor gene deletion library of *aspergillus luchuensis*. *J Gen Appl Microbiol.* 2021;67:118–23.
- Schrevers S, Sanglard D. Hijacking transposable elements for saturation mutagenesis in fungi. *Front Fungal Biol.* 2021;2:1–12.
- Sun M, Gao AX, Ledesma-Amaro R, Li A, Wang R, Nie J, et al. Hypersecretion of OmlA antigen in *Corynebacterium glutamicum* through high-throughput based development process. *Appl Microbiol Biotechnol.* 2022;106:2953–67.
- Celie PHN, Parret AHA, Perrakis A. Recombinant cloning strategies for protein expression. *Curr Opin Struct Biol.* 2016;38:145–54.
- Knight T. Idempotent vector design for standard assembly of BioBricks. 2003.
- Shetty RP, Endy D, Knight TF. Engineering BioBrick vectors from BioBrick parts. *J Biol Eng.* 2008;2:5.
- Shetty R, Lizarazo M, Rettberg R, Knight TF. Assembly of BioBrick standard biological parts using three antibiotic assembly. *Methods Enzymol.* 2011;498:311–26.
- Aubry C, Pernodet J-L, Lautru S. Modular and integrative vectors for synthetic biology applications in *Streptomyces* spp. *Appl Environ Microbiol.* 2019;85:e00485–19.
- Cao Y, Song M, Li F, Li C, Lin X, Chen Y, et al. A synthetic plasmid toolkit for *Shewanella oneidensis* MR-1. *Front Microbiol.* 2019;10:410.
- Wang Y, Gao X, Liu X, Li Y, Sun M, Yang Y, et al. Construction of a 3A system from BioBrick parts for expression of recombinant hirudin variants III in *Corynebacterium glutamicum*. *Appl Microbiol Biotechnol.* 2020;104:8257–66.
- Landy A. Dynamic, structural, and regulatory aspects of lambda site-specific recombination. *Annu Rev Biochem.* 1989;58:913–41.
- Zhang Y, Werling U, Edelmann W. SLiCE: a novel bacterial cell extract-based DNA cloning method. *Nucleic Acids Res.* 2012;40:e55.
- Motohashi K. Seamless ligation cloning extract (SLiCE) method using cell lysates from laboratory *Escherichia coli* strains and its application to slip site-directed mutagenesis. *Methods Mol Biol.* 2017;1498:349–57.
- Okegawa Y, Motohashi K. A simple and ultra-low cost homemade seamless ligation cloning extract (SLiCE) as an alternative to a commercially available seamless DNA cloning kit. *Biochem Biophys Reports.* 2015;4:148–51.
- Aslanidis C, de Jong PJ. Ligation-independent cloning of PCR products (LIC-PCR). *Nucleic Acids Res.* 1990;18:6069–74.
- Eschenfeldt WH, Lucy S, Millard CS, Joachimiak A, Mark ID. A family of LIC vectors for high-throughput cloning and purification of proteins. *Methods Mol Biol.* 2009;498:105–15.
- Stols L, Gu M, Dieckman L, Raffin R, Collart FR, Donnelly MI. A new vector for high-throughput, ligation-independent cloning encoding a tobacco etch virus protease cleavage site. *Protein Expr Purif.* 2002;25:8–15.
- Tachibana A, Tohiguchi K, Ueno T, Setogawa Y, Harada A, Tanabe T. Preparation of long sticky ends for universal ligation-independent cloning: sequential T4 DNA polymerase treatments. *J Biosci Bioeng.* 2009;107:668–9.
- Li MZ, Elledge SJ. Harnessing homologous recombination in vitro to generate recombinant DNA via SLIC. *Nat Methods.* 2007;4:251–6.
- Gibson DG, Young L, Chuang R-Y, Venter JC, Hutchison CA, Smith HO. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods.* 2009;6:343–5.
- Camilo CM, Polikarpov I. High-throughput cloning, expression and purification of glycoside hydrolases using ligation-independent cloning (LIC). *Protein Expr Purif.* 2014;99:35–42.
- Annaluru N, Muller H, Ramalingam S, Kandavelou K, London V, Richardson SM, et al. Assembling DNA fragments by USER fusion. *Methods Mol Biol.* 2012;852:77–95.
- Nour-Eldin HH, Hansen BG, Nørholm MHH, Jensen JK, Halkier BA. Advancing uracil-excision based cloning towards an ideal technique for cloning PCR fragments. *Nucleic Acids Res.* 2006;34:e122.
- Schmid-Burgk JL, Schmidt T, Kaiser V, Höning K, Hornung V. A ligation-independent cloning technique for high-throughput assembly of transcription activator-like effector genes. *Nat Biotechnol.* 2013;31:76–81.
- Yuan H, Peng L, Han Z, Xie J-J, Liu X-P. Recombinant expression library of *Pyrococcus furiosus* constructed by high-throughput cloning: a useful tool for functional and structural genomics. *Front Microbiol.* 2015;6:943.
- Zhang H, Li X, Liu Q, Sun J, Secundo F, Mao X. Construction of a super-folder fluorescent protein-guided secretory expression system for the production of phospholipase D in *Bacillus subtilis*. *J Agric Food Chem.* 2021;69:6842–9.

47. Kovačević G, Ostafe R, Balaž AM, Fischer R, Prodanović R. Development of GFP-based high-throughput screening system for directed evolution of glucose oxidase. *J Biosci Bioeng*. 2019;127:30–7.
48. Kitamura A, Nakayama Y, Kinjo M. Efficient and dynamic nuclear localization of green fluorescent protein via RNA binding. *Biochem Biophys Res Commun*. 2015;463:401–6.
49. Santos-Aberturas J, Dörr M, Waldo GS, Bornscheuer UT. In-depth high-throughput screening of protein engineering libraries by Split-GFP direct crude cell extract data normalization. *Chem Biol*. 2015;22:1406–14.
50. Mo HM, Xu Y, Yu XW. Improved soluble expression and catalytic activity of a thermostable esterase using a high-throughput screening system based on a split-GFP assembly. *J Agric Food Chem*. 2018;66:12756–64.
51. Sun M, Gao AX, Li A, Liu X, Wang R, Yang Y, et al. Bicistronic design as recombinant expression enhancer: characteristics, applications, and structural optimization. *Appl Microbiol Biotechnol*. 2021;105:7709–20.
52. Mendez-Perez D, Gunasekaran S, Orlor VJ, Pfeleger BF. A translation-coupling DNA cassette for monitoring protein translation in *Escherichia coli*. *Metab Eng*. 2012;14:298–305.
53. Gautier A. Fluorescence-activating and absorption-shifting tags for advanced imaging and biosensing. *Acc Chem Res*. 2022;55:3125–35.
54. Balleza E, Kim JM, Cluzel P. Systematic characterization of maturation time of fluorescent proteins in living cells. *Nat Methods*. 2018;15:47–51.
55. Eason MG, Pandelieva AT, Mayer MM, Khan ST, Garcia HG, Chica RA. Genetically encoded fluorescent biosensor for rapid detection of protein expression. *ACS Synth Biol*. 2020;9:2955–63.
56. Grünberger A, Wiechert W, Kohlheyer D. Single-cell microfluidics: opportunity for bioprocess development. *Curr Opin Biotechnol*. 2014;29:15–23.
57. Zhang C, Xing D. Miniaturized PCR chips for nucleic acid amplification and analysis: latest advances and future trends. *Nucleic Acids Res*. 2007;35:4223–37.
58. Weng L, Spoonamore JE. Droplet microfluidics-enabled high-throughput screening for protein engineering. *Micromachines*. 2019;10:734.
59. Jian X, Guo X, Wang J, Tan ZL, Xing X, hui, Wang L, et al. Microbial microdroplet culture system (MMC): an integrated platform for automated, high-throughput microbial cultivation and adaptive evolution. *Biotechnol Bioeng*. 2020;117:1724–37.
60. Bachmann H, Fischlechner M, Rabbers I, Barfa N, Dos Santos FB, Molenaar D, et al. Availability of public goods shapes the evolution of competing metabolic strategies. *Proc Natl Acad Sci U S A*. 2013;110:14302–7.
61. Iftikhar S, Vigne A, Sepulveda-Diaz JE. Droplet-based microfluidics platform for antifungal analysis against filamentous fungi. *Sci Rep*. 2021;11:22998.
62. Barnard GC, Kull AR, Sharkey NS, Shaikh SS, Rittenhour AM, Burnina I, et al. High-throughput screening and selection of yeast cell lines expressing monoclonal antibodies. *J Ind Microbiol Biotechnol*. 2010;37:961–71.
63. Duetz WA. Microtiter plates as mini-bioreactors: miniaturization of fermentation methods. *Trends Microbiol*. 2007;15:469–75.
64. Velez-Suberbie ML, Betts JJP, Walker KL, Robinson C, Zoro B, Keshavarz-Moore E. High throughput automated microbial bioreactor system used for clone selection and rapid scale-down process optimization. *Biotechnol Prog*. 2018;34:58–68.
65. Sandner V, Pybus LP, McCreath G, Glassey J. Scale-down model development in ambr systems: an industrial perspective. *Biotechnol J*. 2019;14:e1700766.
66. Fortney NW, Hanson NJ, Rosa PRF, Donohue TJ, Noguera DR. Diverse profile of fermentation byproducts from thin stillage. *Front Bioeng Biotechnol*. 2021;9:695306.
67. Kaspersetz L, Waldburger S, Schermeyer M-T, Riedel SL, Groß S, Neubauer P, et al. Automated bioprocess feedback operation in a high-throughput facility via the integration of a mobile robotic lab assistant. *Front Chem Eng*. 2022;4:1–13.
68. Ellert A, Grebe A. Process optimization made easy: design of experiments with multi-bioreactor system BIOSTAT® Qplus. *Nat Methods*. 2011;8:i–ii.
69. Moussa M, Ibrahim M, El Ghazaly M, Rohde J, Gnath S, Anton A, et al. Expression of recombinant staphylokinase in the methylotrophic yeast *Hansenula polymorpha*. *BMC Biotechnol*. 2012;12:96.
70. Tan J, Chu J, Shi W, Lin C, Guo Y, Zhuang Y, et al. High-throughput screening strategy used for enhanced production of pigment by *Monascus purpureus* D39-4. *Food Sci Biotechnol*. 2012;21:1603–10.
71. Motta Dos Santos LF, Coutte F, Ravallec R, Dhulster P, Tournier-Couturier L, Jacques P. An improvement of surfactin production by *B. subtilis* ATCC 21332 through the use of a new automatic culture system. *Biotechnol Prog*. 2010;27:1604–9.
72. Hemmerich J, Freier L, Wiechert W, von Lieres E, Oldiges M. Generic protocol for optimization of heterologous protein production using automated micro-bioreactor technology. *J Vis Exp*. 2017;130:56234.
73. Lin C, Lin K, Luong YP, Rao BG, Wei YY, Brennan DL, et al. In vitro resistance studies of hepatitis C virus serine protease inhibitors, VX-950 and BILN 2061: structural analysis indicates different resistance mechanisms. *J Biol Chem*. 2004;279:17508–14.
74. Mandenius C-F, Brundin A. Bioprocess optimization using design-of-experiments methodology. *Biotechnol Prog*. 2008;24:1191–203.
75. Krausch N, Kim JW, Barz T, Lucia S, Groß S, Huber MC, et al. High-throughput screening of optimal process conditions using model predictive control. *Biotechnol Bioeng*. 2022;119:3584–95.
76. Kim JW, Krausch N, Aizpuru J, Barz T, Lucia S, Neubauer P, et al. Model predictive control and moving horizon estimation for adaptive optimal bolus feeding in high-throughput cultivation of *E. coli*. *Comput Chem Eng*. 2023;172:108158.
77. Gill NK, Appleton M, Baganz F, Lye GJ. Design and characterisation of a miniature stirred bioreactor system for parallel microbial fermentations. *Biochem Eng J*. 2008;39:164–76.
78. Janakiraman V, Kwiatkowski C, Kshirsagar R, Ryll T, Huang Y-M. Application of high-throughput mini-bioreactor system for systematic scale-down modeling, process characterization, and control strategy development. *Biotechnol Prog*. 2015;31:1623–32.
79. Fricke J, Pohlmann K, Jonescheit NA, Ellert A, Joksche B, Luttmann R. Designing a fully automated multi-bioreactor plant for fast DoE optimization of pharmaceutical protein production. *Biotechnol J*. 2013;8:738–47.
80. Diaz-Flores E, Meyer T, Giorkallos A. Evolution of artificial intelligence-powered technologies in biomedical research and healthcare. In: Beutel S, Lenk F, editors. *Smart biolabs of the future*. Cham: Springer; 2022. pp. 23–60.
81. Kiss B, Gottschalk U, Pohlscheidt M. *New Bioprocessing strategies: development and manufacturing of recombinant antibodies and proteins*. Springer International Publishing; 2018.
82. Effio CL, Hubbuch J. Next generation vaccines and vectors: Designing downstream processes for recombinant protein-based virus-like particles. *Biotechnol J*. 2015;10:715–27.
83. Dortay H, Akula UM, Westphal C, Sittig M, Mueller-Roeber B. High-throughput protein expression using a combination of ligation-independent cloning (LIC) and infrared fluorescent protein (IFP) detection. *PLoS ONE*. 2011;6:e18900.
84. Heath C, Kiss R. Cell culture process development: advances in process engineering. *Biotechnol Prog*. 2007;23:46–51.
85. Huang C-J, Lin H, Yang X. Industrial production of recombinant therapeutics in *Escherichia coli* and its recent advancements. *J Ind Microbiol Biotechnol*. 2012;39:383–99.
86. Binder D, Drepper T, Jaeger K-E, Delvigne F, Wiechert W, Kohlheyer D, et al. Homogenizing bacterial cell factories: analysis and engineering of phenotypic heterogeneity. *Metab Eng*. 2017;42:145–56.
87. Rienzo M, Jackson SJ, Chao LK, Leaf T, Schmidt TJ, Navidi AH, et al. High-throughput screening for high-efficiency small-molecule biosynthesis. *Metab Eng*. 2021;63:102–25.
88. Markert S, Musmann C, Hülsmann P, Joeris K. Automated and enhanced clone screening using a fully automated microtiter plate-based system for suspension cell culture. *Biotechnol Prog*. 2019;35:e2760.
89. Kondragunta B, Drew JL, Bronson KA, Moreira AR, Rao G. Advances in clone selection using high-throughput bioreactors. *Biotechnol Prog*. 2010;26:1095–103.
90. Krause M, Neubauer A, Neubauer P. The fed-batch principle for the molecular biology lab: controlled nutrient diets in ready-made media improve production of recombinant proteins in *Escherichia coli*. *Microb Cell Fact*. 2016;15:110.
91. Fricke J, Pohlmann K, Tatge F, Lang R, Faber B, Luttmann R. A multi-bioreactor system for optimal production of malaria vaccines with *Pichia pastoris*. *Biotechnol J*. 2011;6:437–51.
92. Harms J, Rathore A. Defining design space for biotech products: Case study of *Pichia pastoris* fermentation. *Biotechnol Prog*. 2010;24:655–62.
93. Shin WS, Lee D, Kim S, Jeong YS, Chun GT. Application of scale-up criterion of constant oxygen mass transfer coefficient ($k_L a$) for production of itaconic acid in a 50 L pilot-scale fermentor by fungal cells of *aspergillus terreus*. *J Microbiol Biotechnol*. 2013;23:1445–53.
94. Thiry M, Cingolani D. Optimizing scale-up fermentation processes. *Trends Biotechnol*. 2002;20:103–5.
95. Kim CH, Rao KJ, Youn DJ, Rhee SK. Scale-up of recombinant hirudin production from *Saccharomyces cerevisiae*. *Biotechnol Bioprocess Eng*. 2003;8:303–5.

96. Schmidt FR. Optimization and scale up of industrial fermentation processes. *Appl Microbiol Biotechnol.* 2005;68:425–35.
97. Junker BH. Scale-up methodologies for *Escherichia coli* and yeast fermentation processes. *J Biosci Bioeng.* 2004;97:347–64.
98. Jungo C, Schenk J, Pasquier M, Marison IW, von Stockar U. A quantitative analysis of the benefits of mixed feeds of sorbitol and methanol for the production of recombinant avidin with *Pichia pastoris*. *J Biotechnol.* 2007;131:57–66.
99. C.Ray S, Behera S, Das R. U, K.Panda S, P.Saranraj. Essentials in fermentation technology. *Learn Mater Biosci.* 2019;313.
100. Holland I, Davies JA. Automation in the life science research laboratory. *Front Bioeng Biotechnol.* 2020;8:1–18.
101. Knepper A, Heiser M, Glauche F, Neubauer P. Robotic platform for parallelized cultivation and monitoring of microbial growth parameters in microwell plates. *SLAS Technol.* 2014;19:593–601.
102. Braun P, LaBaer J. High throughput protein production for functional proteomics. *Trends Biotechnol.* 2003;21:383–8.
103. Khanal SK, Tarafdar A, You S. Artificial intelligence and machine learning for smart bioprocesses. *Bioresour Technol.* 2023;375:128826.
104. Helleckes LM, Hemmerich J, Wiechert W, von Lieres E, Grünberger A. Machine learning in bioprocess development: from promise to practice. *Trends Biotechnol.* 2023;41:817–35.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.